

STIC-Biotech/ChemLib

103886

From: Chan, Christina  
Sent: Monday, September 15, 2003 9:05 AM  
To: Rao, Manjunath N.; STIC-Biotech/ChemLib  
Subject: RE: RUSH sequence search request for 10/083,336

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644  
308-3973  
CM-1, 9B19

-----Original Message-----

From: Rao, Manjunath N.  
Sent: Monday, September 15, 2003 9:00 AM  
To: Chan, Christina  
Subject: RUSH sequence search request for 10/083,336

Hello Christina,

Please authorize the request below as RUSH. The reason being this is an amended due this bi-week and the previous search I did is incomplete.

Thanks  
Manjunath

-----  
From: Manjunath N. Rao  
Art Unit 1652, Room 10A11  
Mail Box in Room 10D 01  
Phone: 306-5681

Date: 9-15-03

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

Please search the following as soon as possible for application with serial number **10/083,336**

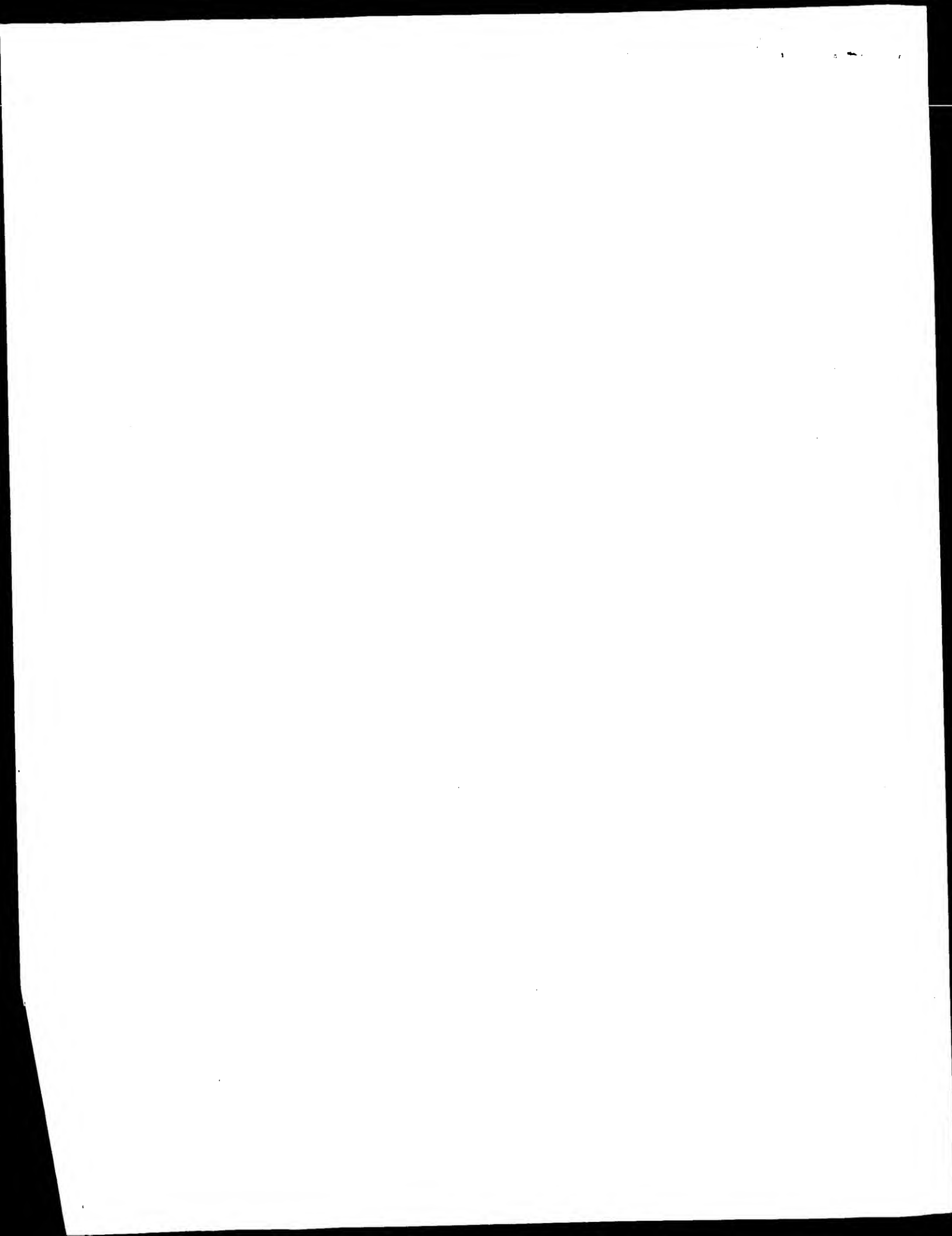
1. SEQ ID NO:1, 2, 3, and 4 against all commercial protein databases including issued patents database and pending application database and provide a print of all results.

If you have any questions please call me at the above phone number.

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 9/16/03  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: 4  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: 102  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_





Thanks

103 886

Manjunath N. Rao, Ph.D.  
Biotechnology Patent Examiner  
Art Unit 1652, Room 10A11  
Mail Box in 10D01  
Crystal Mall 1, USPTO.

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_





# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 103886**

**TO: Manjunath N Rao**  
**Location: cm1/10a11/10d01**  
**Art Unit: 1652**  
**Tuesday, September 16, 2003**

**Case Serial Number: 10/083336**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**CM1-6B02**  
**Phone: 305-9203**

**edward.hart@uspto.gov**

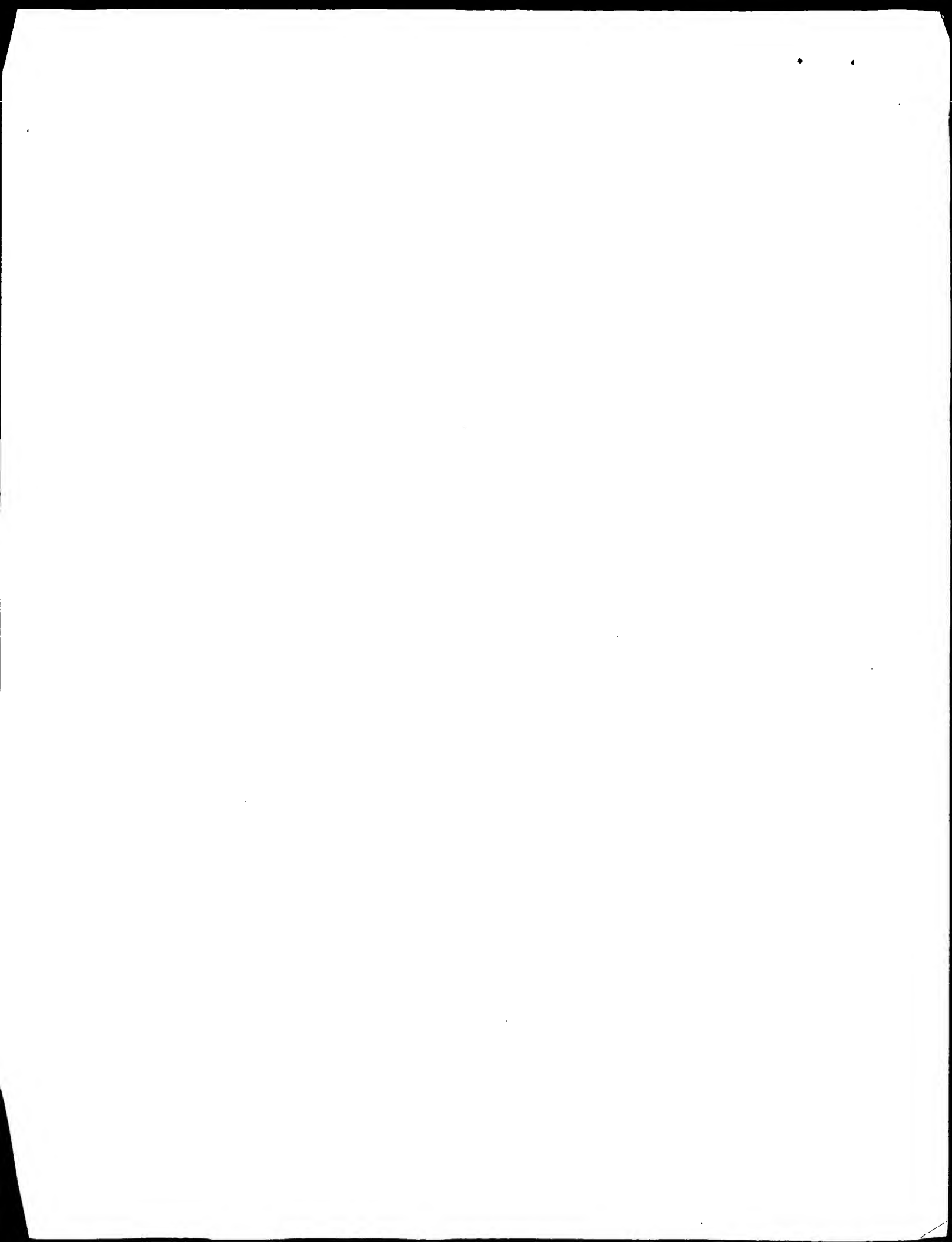
### **Search Notes**

Examiner Rao,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart





# STIC SEARCH RESULTS

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or *contact*:

Mary Hale, Information Branch Supervisor  
308-4258, CM1-1E01

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

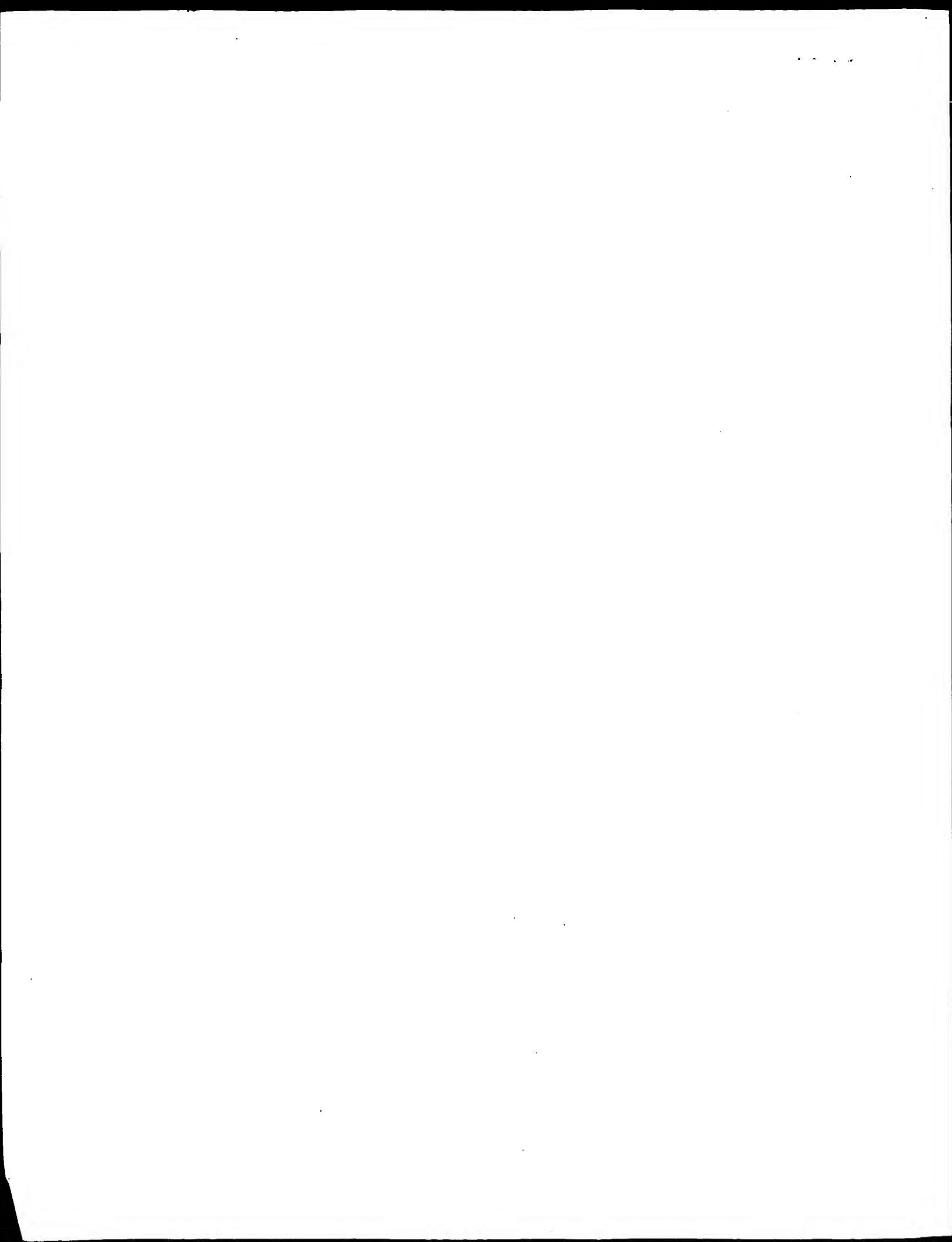
➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 – Circ. Desk





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:29:47 : Search time 95.9159 Seconds

(without alignments)  
953.195 Million cell updates/sec

Title: US-10-083-336a-1

Perfect score: 3051

Sequence: 1 MKPGNTIVIMYAVATWLC.....KQILLYPLHDPNQIMLPUR 576

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
A\_Geneseq\_19Jun03:\*  
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3051	100.0	576	18 AAW25787	Castorbean ricin.
2	3051	100.0	576	20 AAY55892	Castor bean ricin
3	3051	100.0	576	21 AAY78592	Ricinus communis r
4	3051	100.0	576	22 AAG78301	Castor bean prepro
5	3051	100.0	576	22 AAG78302	Castor bean prepro
6	3034	99.4	574	8 AAP70325	Sequence of Ricinu
7	3032	99.4	574	10 AAP94793	DNA sequence of r1
8	2989	98.0	576	8 AAP70326	Sequence of Ricinu
9	2980	97.7	565	6 AAP50166	Sequence of prepro

10	2980	97.7	565	22 AAG78300	Castor bean prepro
11	2968	97.3	565	7 AAP60240	PreproRiclin.
12	2960	97.0	562	10 AAP90079	Ricin D. Ricinus
13	2947	96.6	565	22 AAG78304	Modified castor be
14	2531.5	83.0	540	18 AAW25143	Castor oil plant a
15	2531.5	83.0	540	18 AAW21706	R. communis aglut
16	2480.5	81.3	534	8 AAP70037	Sequence of Ricin
17	1730	56.7	332	8 AAP70097	Ricin A. Escheric
18	1730	56.7	332	8 AAP70838	Sequence of Ricin
19	1730	56.7	332	11 AAP95639	Ricin A encoded by
20	1719	56.3	332	11 AAR06554	Ricin A gene produ
21	1494.5	49.0	290	18 AAW25136	Ricin A-chain ribo
22	1494.5	49.0	290	18 AAW21699	Ricin A-chain RFP
23	1471.5	48.2	534	14 AAR39571	Sequence of G-FIT.
24	1409	46.2	262	9 AAP82896	Ricin B mutlein wit
25	1392	45.6	262	10 AAP90020	Ricin B mutlein wit
26	1377.5	45.1	554	16 AAR70827	B chain of ricin D
27	1372	45.0	267	14 AAR37290	Anti-catact immu
28	1372	45.0	267	16 AAR63902	Ricin A-chain. Un
29	1369	44.9	262	9 AAP82853	Ricin A-chain (RFA
30	1365	44.7	268	14 AAR39570	Ricin B mutlein wit
31	1362	44.6	267	13 AAR30722	Sequence of ricin-
32	1362	44.6	267	21 AAB19265	Ricin A from picil
33	1357	44.5	251	9 AAP81329	Amino acid sequenc
34	1357	44.5	251	10 AAP91887	The ricin B chain
35	1357	44.5	251	10 AAP90837	Amino acid sequenc
36	1356	44.4	262	9 AAP82851	Sequence encoded b
37	1355	44.4	262	9 AAP82852	Ricin B mutlein wit
38	1351	44.3	262	9 AAP82850	Ricin B mutlein wit
39	1345	44.1	267	16 AAP74176	Ricin A chain (RFA
40	1341	44.0	262	9 AAP82849	Ricin B mutlein wit
41	1339	43.9	267	14 AAR32430	Ricin B mutlein wit
42	1334	43.1	280	10 AAP95648	Ricin A. Syntheti
43	1304	42.7	531	20 AAY25879	Ricin agglutinin A
44	1304	42.7	532	20 AAY25882	Mistletoe lectin I
45	1295	42.4	564	18 AAW10021	Mistletoe lectin I

## ALIGNMENTS

RESULT 1	
AAW25787	AAW25787 standard; Protein; 576 AA.
ID	
XX	AAW25787;
AC	
XX	
DT	25-MAR-2003 (updated)
DT	27-MAR-1998 (first entry)
XX	
DE	Castorbean ricin.
XX	
KW	Ricin: cytotoxin; hybrid protein; cell delivery;
KW	cell binding ligand; translocation domain; diptheria toxin B';
KW	Interleukin-2; T-cell lymphoma; organ rejection; therapy.
XX	
OS	Ricinus communis.
XX	
FH	Key
FT	Peptide
FT	1..35
FT	/Label= Sig_peptide
FT	36..302
FT	/Label= A-domain
FT	303..314
FT	/Label= Linker
FT	315..576
FT	/Label= B-domain
XX	
PN	US5668255-A.
XX	
PD	16-SEP-1997.
XX	
PF	04-AUG-1993; 93US-0102387.

```

XX 27-JUN-1991; 91US-0722484.
PR 07-JUN-1984; 84US-0618199.
PR 25-APR-1985; 85US-0726808.
PR 07-JUN-1985; 85US-0742554.
PR 22-DEC-1989; 89US-0456095.
PR 14-JUN-1990; 90US-0538276.
PR 04-AUG-1993; 93US-0102387.
XX (3ERA-) SERAGEN INC.
XX
XX Murphy JR;
XX PI
XX WPI: 1997-470103/43.
DR N-PSDB; AA91638.
XX
XX New hybrid molecules for delivery of agents to cells - comprise a
PT binding domain of a cell binding ligand and a portion of a
PT translocation domain of a protein
XX
XX Example 4; Fig 11A-B; 30pp; English.
XX
XX This polypeptide comprises the castorbean cytotoxin, ricin.
CC DNA (see AA91638) encoding the enzymatic A domain and a portion
CC of the A-to-B linker peptide of ricin was used to construct a
CC ricin-diphtheria toxin B-interleukin-2 gene that was expressed in
CC E. coli. The hybrid protein can be isolated and used to treat
CC conditions involving over-production of cells bearing IL2 receptors,
CC such as certain T-cell lymphomas and organ transplant rejection
CC crises. The hybrid inactivates ribosomes in cells bearing IL2
CC receptors, resulting in cessation of protein synthesis and death of
CC target cells. Claimed hybrid proteins comprise a translocation
CC domain and a cell binding domain from e.g. a hormone, growth factor
CC or polypeptide toxin. The hybrid molecules can be used for the
CC delivery of agents (e.g. therapeutic genes, toxins, detectable
CC labels) into cells. The use of a translocation mechanism ensures
CC that the hybrid will be effective in relatively low doses, since a
CC high proportion of the substance of interest will be taken into the
CC targeted cells. The hybrid molecules can be manufactured as a
CC single hybrid recombinant protein, permitting reproducibility,
CC consistency, and the precise control of composition.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 576 AA:
SQ
Query Match 100.0%; Score 3051; DB 18; Length 576;
Best Local Similarity 100.0%; Pred. No. 1,8e-252;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MKPGNTIYIMYAVATWLCFSGTSGMSFTLEDNNIEPKOPIINFTAGATVQSYTNEI 60
DB 1 MKPGNTIYIMYAVATWLCFSGTSGMSFTLEDNNIEPKOPIINFTAGATVQSYTNEI 60
XX
QY 61 RAVRGLTTGADVRRHEIPVLPNRVGLPIQORFIVELSHAELESTLADVYNAVYGR 120
DB 61 RAVRGLTTGADVRRHEIPVLPNRVGLPIQORFIVELSHAELESTLADVYNAVYGR 120
XX
QY 121 AGNSAFEPHDNOEDAEATLHFTDVQNRVTFAGFNGYRLEBOLGNTLENTELNGPLE 180
DB 121 AGNSAFEPHDNOEDAEATLHFTDVQNRVTFAGFNGYRLEBOLGNTLENTELNGPLE 180
XX
QY 181 EATISALYYSTGGTQLPTLARSFTICIQMISEARFOYIEGEMRTIRYRRSADPSYI 240
DB 181 EATISALYYSTGGTQLPTLARSFTICIQMISEARFOYIEGEMRTIRYRRSADPSYI 240
XX
QY 241 TLENSWGRISTALQESNOGAFASPTOLORRNSKFSYVYVSTILPIIAMVRCAPPPSS 300
DB 241 TLENSWGRISTALQESNOGAFASPTOLORRNSKFSYVYVSTILPIIAMVRCAPPPSS 300
XX
QY 301 OFSLILRPVYVNFNADVCMDEPIYIVGRNGICVDYRDGRPHNGNAIOLMPCCKSNTDAN 360
DB 301 OFSLILRPVYVNFNADVCMDEPIYIVGRNGICVDYRDGRPHNGNAIOLMPCCKSNTDAN 360

```

```

QY 361 QLWTLKRDNTIRSGKLTFTYGYSPGVYIMYDCNTAATADATRMQIMDNGTIINRSSLY 420
DB 361 QLWTLKRDNTIRSGKLTFTYGYSPGVYIMYDCNTAATADATRMQIMDNGTIINRSSLY 420
XX
QY 421 LAATSGNSGTTLTVOGNIIYAVSOGWLPYNNQPFVTTYGLGLQIANSQGVWIEDCS 480
DB 421 LAATSGNSGTTLTVOGNIIYAVSOGWLPYNNQPFVTTYGLGLQIANSQGVWIEDCS 480
XX
QY 481 EKADQOMALYADGSTRPOONRDCLTSDSNIRETVYKILSCGPASSGQRMFKNDGTILN 540
DB 481 EKADQOMALYADGSTRPOONRDCLTSDSNIRETVYKILSCGPASSGQRMFKNDGTILN 540
XX
QY 541 LYSGLVLDVVRASDPSLKOIILPLHGDPMQIWLPLF 576
DB 541 LYSGLVLDVVRASDPSLKOIILPLHGDPMQIWLPLF 576
XX
RESULT 2
AA55892
ID AA55892 standard; Protein; 576 AA.
XX
XX AA55892;
XX
XX 15-FEB-2000 (first entry)
XX
XX Castor bean ricin toxin.
XX
XX Recombinant: hybrid; binding domain; ligand; animal cell; diphtheria;
XX translocation domain; botulinum; neurotoxin; ricin; cholera; tetanus;
XX shiga-like toxin; pertussis; translocation; cytoplasmic membrane; HIV;
XX cytosol; therapy; genetic deficiency disease; enzyme; co-factor; poison;
XX adipocyte; cancer; virus; infection; antibody.
XX
XX Ricinus communis.
XX
XX US9565406-A.
XX
XX 12-OCR-1999.
XX
XX 07-JUN-1995; 95US-0488246.
XX
XX 04-AUG-1993; 93US-0102387.
XX 07-JUN-1984; 84US-0618199.
XX 27-JUN-1991; 91US-0722484.
XX 25-APR-1985; 85US-0726808.
XX 07-JUN-1985; 85US-0742554.
XX 22-DEC-1989; 89US-0456095.
XX 14-JUN-1990; 90US-0538276.
XX
XX (3ERA-) SERAGEN INC.
XX
XX Murphy JR;
XX WPI: 1999-632431/54.
XX N-PSDB; AA230663.
XX
XX Recombinant DNA molecule encoding a three part hybrid protein used in
XX the treatment of AIDS and genetic deficiency diseases -
XX
XX Example 4; Fig 11; 31pp; English.
XX
XX The invention relates to a recombinant DNA molecule encoding a hybrid
XX protein comprising three parts: (a) the first part comprises a portion
XX of the binding domain of a cell-binding polypeptide ligand allowing the
XX hybrid protein to bind to an animal cell; (b) the second part comprises
XX a portion of a translocation domain of a naturally occurring protein
XX selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera
XX toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus
XX toxin, which translocate the third part of the across the cytoplasmic
XX membrane into the cytosol of the cell; and (c) the third part comprises
XX a polypeptide entity to be introduced into the cell, which is non-native
XX to the naturally occurring protein of (b). This sequence represents the
XX Castor bean ricin toxin sequence for use in generating the hybrid of the

```



CC invention. The hybrid molecule enables the direction of appropriate  
CC therapy to affected cells, allowing them to function properly and  
CC alleviate or cure the disease. The hybrid is especially used in treating  
CC genetic deficiency diseases, by delivering to affected cells an enzyme  
CC supplying the missing function, to supplementing cellular levels of a  
CC particular enzyme or a scarce precursor or cofactor, to directing toxins  
CC or other poisons to destroy particular cells (such as adipocytes, cancer  
CC cell, or virus infected cells), to counteracting viral infections such as  
CC HIV, by introducing appropriate antibodies to viral proteins. It is also  
CC involved in the process of getting non-therapeutic substances such as  
CC detectable labels into cells.

CC Sequence 576 AA:

Query Match 100.0%; Score 3051; DB 20; Length 576;  
Best Local Similarity 100.0%; Pred. No. 1.8e-252;  
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPGNTIYIMVAVATWLCFSGSTGMSFTLEDNNIFPKOYPIINFTAGATVQSYNFI 60  
DB 1 MKPGNTIYIMVAVATWLCFSGSTGMSFTLEDNNIFPKOYPIINFTAGATVQSYNFI 60  
QY 61 RAVRGRLTTGADVREHPIVLPNRYGLPINQRFILVELSNHAEISVTLADVTNAYVGYR 120  
DB 61 RAVRGRLTTGADVREHPIVLPNRYGLPINQRFILVELSNHAEISVTLADVTNAYVGYR 120  
QY 121 AGNSAFYFHPDNOEDAEITLFTDVQNRRTFAFGGNYRDLBOLAGNLENIELGNPLE 180  
DB 121 AGNSAFYFHPDNOEDAEITLFTDVQNRRTFAFGGNYRDLBOLAGNLENIELGNPLE 180  
QY 181 EASIALYVSTGGTQPTLARSFIICIQMISEARPOYLEGEMRTIRYRRSADPSYI 240  
DB 181 EASIALYVSTGGTQPTLARSFIICIQMISEARPOYLEGEMRTIRYRRSADPSYI 240  
QY 241 TLENSGRLSTAQESNOCASFPIQLOLRNSKESVYVSLIPIALMYRCAPPPSS 300  
DB 241 TLENSGRLSTAQESNOCASFPIQLOLRNSKESVYVSLIPIALMYRCAPPPSS 300  
QY 301 QESLLRPVYPNNAVCMDEPIYIVIRGNGLCVYRDGRFHNGNAIOLMPCSKSTIDAN 360  
DB 301 QESLLRPVYPNNAVCMDEPIYIVIRGNGLCVYRDGRFHNGNAIOLMPCSKSTIDAN 360  
QY 361 QLWTLKRDNTIRNSGKCLFTTGYSPGVYMIYDCNTAATADATRMQIMDNGTIINRSSLY 420  
DB 361 QLWTLKRDNTIRNSGKCLFTTGYSPGVYMIYDCNTAATADATRMQIMDNGTIINRSSLY 420  
QY 421 LAATSGNSGTLTVQNNIYAVSOGMLPTNMQPPTTYIGLYGLCLQANSQGVWIEDCS 480  
DB 421 LAATSGNSGTLTVQNNIYAVSOGMLPTNMQPPTTYIGLYGLCLQANSQGVWIEDCS 480  
QY 481 EKAQGMALYADDSIRPOQNRDNCILSDSNIREFYVKILSCGPASSGQRMKNGTILN 540  
DB 481 EKAQGMALYADDSIRPOQNRDNCILSDSNIREFYVKILSCGPASSGQRMKNGTILN 540  
QY 541 LYSGLVLDVYRASDPSLKOITLLYPLHGPQNIWLPLE 576  
DB 541 LYSGLVLDVYRASDPSLKOITLLYPLHGPQNIWLPLE 576

RESULT 3  
AA78592  
ID AA78592 standard; Protein; 576 AA.

XX AC AA78592;

XX 05-MAY-2000 (first entry)

XX Rictinus communis rictin protein sequence.

KM Rictin: toxin; hybrid protein; translocation domain; cell destruction;  
KM cell binding domain; genetic deficiency disease; cell targeting; cancer;  
KM adipocyte; enzyme delivery; anti-viral; HIV.

OS Rictinus communis.

XX US6022950-A.

XX 08-FEB-2000.

XX 07-JUN-1995; 9505-0479510.

XX 07-JUN-1984; 84US-0618199.

XX 27-JUN-1991; 91US-0722484.

XX 25-APR-1985; 85US-0726808.

XX 07-JUN-1985; 85US-0742534.

XX 22-DEC-1989; 89US-0456095.

XX 14-JUN-1990; 90US-0538276.

XX 04-AUG-1993; 93US-0102387.

XX (SEEA-) SERAGEN INC.

XX Murphy JR;

XX WPI; 2000-160390/14.

XX N-PSDB; AA290019.

XX Example 4; Fig 11; 32pp; English.

This sequence represents the Rictinus communis rictin protein sequence. The  
CC toxin can be included in the hybrid protein of the invention and used to  
CC destroy or modify the cell that the hybrid protein is targeted to. The  
CC hybrid protein comprises a first part which is a portion of the binding  
CC domain of a cell-binding ligand, effective to cause the hybrid molecule  
CC to bind to a cell of an animal. The second part comprises a portion of a  
CC translocation domain of a naturally occurring protein (e.g. the  
CC translocation domain of diphtheria toxin) the second part translocates  
CC the third part across the cytoplasmic membrane and into the cytosol of  
CC the cell. The third part comprises a chemical entity to be introduced  
CC into the cell, where each of the first and third part is non-native with  
CC respect to naturally occurring protein, and the covalent bond attaching  
CC the second and third part is cleavable. The toxin represented by the  
CC present sequence can form part of the third portion of the hybrid  
CC protein. The cell binding domain binds to a specific cell and the  
CC translocation domain transfers the hybrid molecule across the cell  
CC membrane into the cytosol. The third part of the protein, linked to the  
CC function. The hybrid molecules are useful for treating genetic deficiency  
CC diseases by delivering to affected cells an enzyme supplying the missing  
CC function, to supplement cellular levels of a particular enzyme or a  
CC scarce precursor or cofactor, to direct toxins or other poisons to  
CC destroy particular cells (such as adipocytes, cancer cells, or  
CC virus-infected cells), and to counteract viral infections such as HIV by  
CC introducing into appropriate cells antibodies to viral proteins.

SO Sequence 576 AA:

Query Match 100.0%; Score 3051; DB 21; Length 576;  
Best Local Similarity 100.0%; Pred. No. 1.8e-252;  
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPGNTIYIMVAVATWLCFSGSTGMSFTLEDNNIFPKOYPIINFTAGATVQSYNFI 60  
DB 1 MKPGNTIYIMVAVATWLCFSGSTGMSFTLEDNNIFPKOYPIINFTAGATVQSYNFI 60  
QY 61 RAVRGRLTTGADVREHPIVLPNRYGLPINQRFILVELSNHAEISVTLADVTNAYVGYR 120  
DB 61 RAVRGRLTTGADVREHPIVLPNRYGLPINQRFILVELSNHAEISVTLADVTNAYVGYR 120  
QY 121 AGNSAFYFHPDNOEDAEITLFTDVQNRRTFAFGGNYRDLBOLAGNLENIELGNPLE 180  
DB 121 AGNSAFYFHPDNOEDAEITLFTDVQNRRTFAFGGNYRDLBOLAGNLENIELGNPLE 180

```

QY 181 EAISALYYSTGTGTOPLTLARSFITCIOMISEARFOYIEGEMRTIRYNRRSADPPSVI 240
    |||||||
Db 181 EAISALYYSTGTGTOPLTLARSFITCIOMISEARFOYIEGEMRTIRYNRRSADPPSVI 240
QY 241 TLENSMGRSLTAIOESNOGAFASPIQLORRNGSKFSYDVSTILPIALMYRCAPPPSS 300
    |||||||
Db 241 TLENSMGRSLTAIOESNOGAFASPIQLORRNGSKFSYDVSTILPIALMYRCAPPPSS 300
QY 301 QESLLIRPVVFNPNADVCMDEPIYRIVGRNGLCVDVDRGRHNGNAIQOLMPCKSNMTAN 360
    |||||||
Db 301 QESLLIRPVVFNPNADVCMDEPIYRIVGRNGLCVDVDRGRHNGNAIQOLMPCKSNMTAN 360
QY 361 QLMTLKRDNTIRSNKCLTLYGSPGVYVMIYDCNTAATDAATRMQIMDNGTIINRSSLY 420
    |||||||
Db 361 QLMTLKRDNTIRSNKCLTLYGSPGVYVMIYDCNTAATDAATRMQIMDNGTIINRSSLY 420
QY 421 LAATSGNSGTTLTVOFTNIYASOGWLPNTNTOPEFTTVIGLYGLCLQANSQGVWIEDCSS 480
    |||||||
Db 421 LAATSGNSGTTLTVOFTNIYASOGWLPNTNTOPEFTTVIGLYGLCLQANSQGVWIEDCSS 480
QY 481 EKAEOQWALYADGSIIRPOQRNDCLTSDSNIRETVYVKILSCGPASSGGRMMFKNDGTILN 540
    |||||||
Db 481 EKAEOQWALYADGSIIRPOQRNDCLTSDSNIRETVYVKILSCGPASSGGRMMFKNDGTILN 540
QY 541 LYSGLVLDVRASDPSLKQIILYPLHGDPPQIWLPLF 576
    |||||||
Db 541 LYSGLVLDVRASDPSLKQIILYPLHGDPPQIWLPLF 576

RESULT 4
AAG78301
ID AAG78301 standard; Protein: 576 AA.
XX
AC AAG78301;
XX
DT 15-NOV-2001 (first entry)
XX
DE Castor bean preprotricin protein (SEQ ID 2).
XX
KW Castor bean plant; preprotricin; ricin; A chain; B chain;
KM human immunodeficiency virus infection; HIV; toxin; antiviral agent;
KW retroviral infection; anti-HIV; virucide activity; viral protease.
XX
OS Ricinus communis.
XX
FH Key Location/Qualifiers
FT Peptide 1..35
FT Protein /label= Signal peptide
FT Protein /label= 36..302
FT Protein /label= Ricin_A_chain
FT Peptide /note= "N-glycosidase"
FT Peptide 303..314
FT Peptide /label= linker peptide
FT Protein /note= "Cleaved during activation of ricin"
FT Protein 315..576
FT Protein /label= Ricin_B_chain
FT Protein /note= "Galactose/N-acetylglucosamine-binding lectin"

WC200160393-A1.
XX
PD 23-AUG-2001.
XX
PF 15-FEB-2001; 2001WO-US05282.
XX
PR 15-FEB-2000; 2000US-0182759.
XX
PA (BECH-) BECHTEL BWXT IDAHO LLC.
XX
PI Kaener WK, Ward TE;
XX
DR WPI: 2001-581908/65.
DR N-PSDB; AA164138.
XX

```

```

PT Novel composition comprising toxin e.g., ricin based antiviral compound
PT useful for treating viral infections such as human immunodeficiency
PT virus infection.
XX
XX Disclosure; Page 50-54; 66pp; English.
XX
PS The sequence relates to preprotricin protein encoded by the DNA sequence
CC given in AA164138. The invention relates to a novel toxin (e.g., ricin)
CC based antiviral agent which is toxic to virus-infected cells, but
CC non-toxic to uninfected cells. The invention has anti-HIV and virucide
CC activities. Its mechanism of action is through inactivation of cellular
CC ribosomes and enhancement of binding of the antiviral agent to galactose
CC residues on cell surfaces, and its cellular internalisation. The
CC invention is useful for treating human immunodeficiency virus infection
CC and other viral infections, especially retroviral infections. The
CC antiviral agent is activated in viral particles or early-stage infected
CC cells, killing the cells upon infection and effectively preventing the
CC integration of the viral genome into the host genome thereby preventing
CC the latency/rebound problem. The agent enters all HIV susceptible cells,
CC and not just cells known to act as host cells for the virus. The
CC antiviral agent remains inert in a cell until degraded in it, unless the
CC cell is infected with the virus, where the viral protease activates it.
XX
SQ Sequence 576 AA:
Query Match 100.0%; Score 3051; DB 22; Length 576;
Best Local Similarity 100.0%; Pred. No. 1,8e-252;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKPGNTIYIMTAVATWLCFGSTSGNSFTLEDNNIFPKOYPIINFTAGATVOSTNFI 60
Db 1 MKPGNTIYIMTAVATWLCFGSTSGNSFTLEDNNIFPKOYPIINFTAGATVOSTNFI 60
QY 61 RAVGRLLTGADVRHEHPVLPNVRGLPINORFLVLSNHAELSTVALDVNAYVGR 120
Db 61 RAVGRLLTGADVRHEHPVLPNVRGLPINORFLVLSNHAELSTVALDVNAYVGR 120
QY 121 AGNSAYFEHFDNOEDAETHLFTDVQNRRTFAFGNGYRLRLQALAGNLENIELGNGPLE 180
Db 121 AGNSAYFEHFDNOEDAETHLFTDVQNRRTFAFGNGYRLRLQALAGNLENIELGNGPLE 180
QY 181 EAISALYYSTGTGTOPLTLARSFITCIOMISEARFOYIEGEMRTIRYNRRSADPPSVI 240
Db 181 EAISALYYSTGTGTOPLTLARSFITCIOMISEARFOYIEGEMRTIRYNRRSADPPSVI 240
QY 241 TLENSMGRSLTAIOESNOGAFASPIQLORRNGSKFSYDVSTILPIALMYRCAPPPSS 300
Db 241 TLENSMGRSLTAIOESNOGAFASPIQLORRNGSKFSYDVSTILPIALMYRCAPPPSS 300
QY 301 QESLLIRPVVFNPNADVCMDEPIYRIVGRNGLCVDVDRGRHNGNAIQOLMPCKSNMTAN 360
Db 301 QESLLIRPVVFNPNADVCMDEPIYRIVGRNGLCVDVDRGRHNGNAIQOLMPCKSNMTAN 360
QY 361 QLMTLKRDNTIRSNKCLTLYGSPGVYVMIYDCNTAATDAATRMQIMDNGTIINRSSLY 420
Db 361 QLMTLKRDNTIRSNKCLTLYGSPGVYVMIYDCNTAATDAATRMQIMDNGTIINRSSLY 420
QY 421 LAATSGNSGTTLTVOFTNIYASOGWLPNTNTOPEFTTVIGLYGLCLQANSQGVWIEDCSS 480
Db 421 LAATSGNSGTTLTVOFTNIYASOGWLPNTNTOPEFTTVIGLYGLCLQANSQGVWIEDCSS 480
QY 481 EKAEOQWALYADGSIIRPOQRNDCLTSDSNIRETVYVKILSCGPASSGGRMMFKNDGTILN 540
Db 481 EKAEOQWALYADGSIIRPOQRNDCLTSDSNIRETVYVKILSCGPASSGGRMMFKNDGTILN 540
QY 541 LYSGLVLDVRASDPSLKQIILYPLHGDPPQIWLPLF 576
Db 541 LYSGLVLDVRASDPSLKQIILYPLHGDPPQIWLPLF 576

RESULT 5
AAG78302
ID AAG78302 standard; Protein: 576 AA.

```



XX  
PIatak M;

XX  
DR WPI: 1987-265177/38.  
XX  
DR N-PSDB: AAN70525.

PT New non-glycosylated ricin precursor and toxin etc. - are prepd.  
PT by recombinant DNA procedures with specific isolation steps for  
PT purer and soluble prods.

XX  
PS Disclosure: Fig 13(1-2); 112pp; English.

XX  
CC The full-length sequences encoding ricin A (AAN70520), ricin D  
CC (AAN70525), putative ricin E (AAN70526) and RCA (AAN70524) in precursor  
CC form were obtd. using messenger RNA to obtain a cDNA library, and  
CC then probing the library to retrieve the desired cDNA inserts. The  
CC library was probed using the 35-mer given in AAN70514. Figure 4 (see  
CC AAN70520, AAN70521, AAN70522) shows the nucleotide sequences of three  
CC plasmids contg. cDNA inserts obtd. by probing a cDNA library for  
CC sequences encoding ricin B using the probe in AAN70517. The cDNA  
CC inserts can be placed into expression vectors. Site-directed  
CC mutagenesis may be used to place an ARG start codon and a HindIII  
CC site at the beginning of the mature protein (see AAN70518). The  
CC coding sequences of the inserts can be ligated into expression  
CC vectors contg. the Phoa promoter-operator and leader sequence  
CC (AAN70523) and suitable retroregulators.  
CC (Updated on 25-MAR-2003 to correct PA field.)

XX  
SQ Sequence 574 AA;

Query Match 99.4%; Score 3034; DB 8; Length 574;

Best Local Similarity 99.8%; Pred. No. 5.2e-251;

Matches 573; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PGGNTIVIMYAVATWLCFGSTSGWSTLEDNNIFPKOYPIINFTTAGATVQSYTNFIRA 62  
DB 1 PGGNTIVIMYAVATWLCFGSTSGWSTLEDNNIFPKOYPIINFTTAGATVQSYTNFIRA 60  
QY 63 VRGRITGADVREHIEPVLPNRYGLPINOREFLVELSNHAEISVTLADVTNAYVGYRAG 122  
DB 61 VRGRITGADVREHIEPVLPNRYGLPINOREFLVELSNHAEISVTLADVTNAYVGYRAG 120  
QY 123 NSAYFFHDPNOEDAEATHTLFTDVONRYTFAFGNRYDLRLEQAGNLENIELGNPLEEA 182  
DB 121 NSAYFFHDPNOEDAEATHTLFTDVONRYTFAFGNRYDLRLEQAGNLENIELGNPLEEA 180  
QY 183 ISALYYSTGCTQLPTLARSFTICIQMISEARFOYIEGEMRTIRYNNRSAPDPSSVITL 242  
DB 181 ISALYYSTGCTQLPTLARSFTICIQMISEARFOYIEGEMRTIRYNNRSAPDPSSVITL 240  
QY 243 ENSMGRSLTAIOESNOGAFASPIQLORRNGSKFSYDVSLIPITIAMVYRCAPPSSQF 302  
DB 241 ENSMGRSLTAIOESNOGAFASPIQLORRNGSKFSYDVSLIPITIAMVYRCAPPSSQF 300  
QY 303 SLILRPVVPNFNADVCMDEPIVIRVGNGLCVDRGDRFHNGNAIOLWPKSNTDANQL 362  
DB 301 SLILRPVVPNFNADVCMDEPIVIRVGNGLCVDRGDRFHNGNAIOLWPKSNTDANQL 360  
QY 363 WTIKRDNTIRNGKCTITYGSPGVYVYIYDCNTAANDPRTMOWIMDGTIINPSSVILA 422  
DB 361 WTIKRDNTIRNGKCTITYGSPGVYVYIYDCNTAANDPRTMOWIMDGTIINPSSVILA 420  
QY 423 ATSGNSGTTLTIVOTNIYAVSOGMLPTNNTQPEFTTIVGLVLCLOANSQGVWIEDCSSEK 482  
DB 421 ATSGNSGTTLTIVOTNIYAVSOGMLPTNNTQPEFTTIVGLVLCLOANSQGVWIEDCSSEK 480  
QY 483 AEQOQALYADGSIRPOONRDNCLTSDSNIRETVYKIIISCGPASSGORMEKNDGTIINLY 542  
DB 481 AEQOQALYADGSIRPOONRDNCLTSDSNIRETVYKIIISCGPASSGORMEKNDGTIINLY 540  
QY 543 SGLVLDVRAVSPSLKQIITLYPLHGPNOIWLPLF 576  
DB 541 SGLVLDVRAVSPSLKQIITLYPLHGPNOIWLPLF 574

RESULT 7

AAP94793  
ID AAP94793 standard; protein; 574 AA.

XX  
AC AAP94793;

XX  
DT 06-JUL-1990 (first entry)

DE DNA sequence of ricin toxin D.

XX  
KM Ricin; toxin; baculovirus; mutants; ss.

XX  
OS Baculovirus.

XX  
FH Key Location/Qualifiers

FT Reptide 1..35 /Label=Leader peptide.

FT Domain 36..300 /Label=A-chain

FT Domain 301..574 /Label=B-chain

XX  
PN WO8901037-A.

XX  
PD 09-FEB-1989.

XX  
PE 20-JUL-1988; 88WO-US02442.

XX  
PR 24-JUL-1987; 87US-0077126.

XX  
PR 08-FEB-1988; 88US-0153778.

XX  
PA (CETU ) CETUS CORP.

XX  
PI Houston LL, Lane JA, Piatak M, Clark R;

XX  
DR WPI: 1989-061173/08.

XX  
DR N-PSDB: AAN91039.

PT Recombinant baculovirus transfer vectors -

PT used for prodn. of ricin toxin in a baculovirus insect cell

PT expression system.

XX  
PS Disclosure: Page -: 85pp; English.

CC Recombinant baculovirus vector is capable of introducing ricin gene into

CC virus genome, allowing creation of a insect cell/baculovirus ricin

CC expression system.

XX  
SQ Sequence 574 AA;

Query Match 99.4%; Score 3032; DB 10; Length 574;

Best Local Similarity 99.8%; Pred. No. 7.7e-251;

Matches 573; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PGGNTIVIMYAVATWLCFGSTSGWSTLEDNNIFPKOYPIINFTTAGATVQSYTNFIRA 62  
DB 1 PGGNTIVIMYAVATWLCFGSTSGWSTLEDNNIFPKOYPIINFTTAGATVQSYTNFIRA 60  
QY 63 VRGRITGADVREHIEPVLPNRYGLPINOREFLVELSNHAEISVTLADVTNAYVGYRAG 122  
DB 61 VRGRITGADVREHIEPVLPNRYGLPINOREFLVELSNHAEISVTLADVTNAYVGYRAG 120  
QY 123 NSAYFFHDPNOEDAEATHTLFTDVONRYTFAFGNRYDLRLEQAGNLENIELGNPLEEA 182  
DB 121 NSAYFFHDPNOEDAEATHTLFTDVONRYTFAFGNRYDLRLEQAGNLENIELGNPLEEA 180  
QY 183 ISALYYSTGCTQLPTLARSFTICIQMISEARFOYIEGEMRTIRYNNRSAPDPSSVITL 242  
DB 181 ISALYYSTGCTQLPTLARSFTICIQMISEARFOYIEGEMRTIRYNNRSAPDPSSVITL 240  
QY 243 ENSMGRSLTAIOESNOGAFASPIQLORRNGSKFSYDVSLIPITIAMVYRCAPPSSQF 302

```

|||||
Db 241 ENSWGRLSTAIQESNCGAFASPIQLQRRNGSKFSYDVSLIPITLAMYRCAPPPSSQF 300
Qy 303 SLILIRPVVFNADVCDPEPIRYIRGRNLCVDVDRGRHNGNAIQLPCKSNTDANOL 362
Db 301 SLILIRPVVFNADVCDPEPIRYIRGRNLCVDVDRGRHNGNAIQLPCKSNTDANOL 360
Qy 363 WTLKRDNTIRSNKCLTTTGYSPGYVYMIYDCNTAATDARMOIWDNGTINPRSSLYLA 422
Db 361 WTLKRDNTIRSNKCLTTTGYSPGYVYMIYDCNTAATDARMOIWDNGTINPRSSLYLA 420
Qy 423 ATSGNSGTTLTVOYTNIAVSQGLPTNNTOPEVTTIVGLYGLCLQANSQGVWIEDCSSEK 482
Db 421 ATSGNSGTTLTVOYTNIAVSQGLPTNNTOPEVTTIVGLYGLCLQANSQGVWIEDCSSEK 480
Qy 483 AEOQOMALYADGSIIRPOONRDNCLTSDSNIRETVYKILSCGPASSGQRMFKNDGTTILNY 542
Db 481 AEOQOMALYADGSIIRPOONRDNCLTSDSNIRETVYKILSCGPASSGQRMFKNDGTTILNY 540
Qy 543 SGLVLDVRASDPSLKOIILYPLHGDNPQIWLPLF 576
Db 541 SGLVLDVRASDPSLKOIILYPLHGDNPQIWLPLF 574

RESULT 8
AAP70326
ID AAP70326 standard; Protein: 576 AA.
XX
AC AAP70326;
XX
DT 25-MAR-2003 (updated)
DT 21-MAY-1991 (first entry)
XX
DE Sequence of Ricinus communis (castor bean) Ricin toxin
DE (RT or ricin) E precursor encoded by pr138.
XX
KM Lectin; toxin protein; cytotoxic; cytostatic; castor bean;
KM plant toxin.
XX
OS Ricinus communis.
XX
FH Key Location/Qualifiers
FH Region 1..35
FH /note= "leader"
FH Region 36..302
FH /note= "A-chain"
FH Region 315..576
FH /note= "B-chain"
XX
PN EP237676-A.
XX
PD 23-SEP-1987.
XX
PF 13-NOV-1986; 86EP-0308877.
XX
PR 07-MAR-1986; 86US-0837583.
XX
PA (CETU ) CETUS CORP.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Platak M;
XX
DR WPI: 1987-265177/38.
DR N-PSDB: AAN70526.
XX
PT New non-glycosylated ricin precursor and toxin etc. - are prepd.
PT by recombinant DNA procedures with specific isolation steps for
XX purer and soluble prods.
PS Disclosure: Fig 14(1-2); 112pp; English.
XX
CC The full length sequences encoding ricin A (AAN70520), ricin D
CC (AAN70525) putative ricin E (AAN70526) and RCA (AAN70524) in precursor

```

```

CC form were obtained, using the messenger RNA to obtain a cDNA library, and
CC then probing the library to retrieve the desired cDNA inserts. The
CC library was probed using the 35-mer given in AAN70514. Figure 4 (see
CC AAN70520, AAN70521, AAN70522), shows the nucleotide sequences of three
CC plasmids containing cDNA inserts obtained by probing a cDNA library
CC for sequences encoding ricin B using the probe in AAN70517. The cDNA
CC inserts can be placed into expression vectors. Site-directed
CC mutagenesis may be used to place an ATG start codon and a HindIII
CC site at the beginning of the mature protein. (see AAN70518). The
CC coding sequences of the inserts can be ligated into expression
CC vectors containing the Phoa promoter-operator and leader sequence
CC (AAN70523) and suitable retroregulators.
CC (Updated on 25-MAR-2003 to correct pa field.)
XX

SQ Sequence 576 AA:
Query Match 98.0%; Score 2989; DB 8; Length 576;
Best Local Similarity 97.4%; Pred. No. 3.7e-247;
Matches 561; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKPGGNTIYIMYAVATWLCFGSTSGMSFTLEDNNIPKQYPIINFATTAGATVOSYTNFI 60
Db 1 MKPGGNTIYIMYAVATWLCFGSTSGMSFTLEDNNIPKQYPIINFATTAGATVOSYTNFI 60
Qy 61 RAVRGRLTTGADVREHIEIVLPNRYGLPIINORFTIIVELSNHAEISVTLALDVTNAYVYGYR 120
Db 61 RAVRGRLTTGADVREHIEIVLPNRYGLPIINORFTIIVELSNHAEISVTLALDVTNAYVYGYR 120
Qy 121 AGNSAYFFHPDNOEDAEATHLFTDVONRYTFAFGGNYDRLEQLAGNLRENIEIGNGPLE 180
Db 121 AGNSAYFFHPDNOEDAEATHLFTDVONRYTFAFGGNYDRLEQLAGNLRENIEIGNGPLE 180
Qy 181 EALSALYYSTGGQPLPLARSFTIICIMISEARFOYIEEMTRIRYNRRSAPDSYI 240
Db 181 EALSALYYSTGGQPLPLARSFTIICIMISEARFOYIEEMTRIRYNRRSAPDSYI 240
Qy 241 TLENSWGRLSTAIOESNCGAFASPIQLQRRNGSKFSYDVSLIPITLAMYRCAPPPSS 300
Db 241 TLENSWGRLSTAIOESNCGAFASPIQLQRRNGSKFSYDVSLIPITLAMYRCAPPPSS 300
Qy 301 QFSLILIRPVVFNADVCDPEPIRYIRGRNLCVDVDRGRHNGNAIQLPCKSNTDAN 360
Db 301 QFSLILIRPVVFNADVCDPEPIRYIRGRNLCVDVDRGRHNGNAIQLPCKSNTDAN 360
Qy 361 QLWTLKRDNTIRSNKCLTTTGYSPGYVYMIYDCNTAATDARMOIWDNGTINPRSSLY 420
Db 361 QLWTLKRDNTIRSNKCLTTTGYSPGYVYMIYDCNTAATDARMOIWDNGTINPRSSLY 420
Qy 421 LAATSGNSGTTLTVOYTNIAVSQGLPTNNTOPEVTTIVGLYGLCLQANSQGVWIEDCSSEK 480
Db 421 LAATSGNSGTTLTVOYTNIAVSQGLPTNNTOPEVTTIVGLYGLCLQANSQGVWIEDCSSEK 480
Qy 481 EKAEOQOMALYADGSIIRPOONRDNCLTSDSNIRETVYKILSCGPASSGQRMFKNDGTTILN 540
Db 481 EKAEOQOMALYADGSIIRPOONRDNCLTSDSNIRETVYKILSCGPASSGQRMFKNDGTTILN 540
Qy 541 LYSGLVLDVRASDPSLKOIILYPLHGDNPQIWLPLF 576
Db 541 LYSGLVLDVRASDPSLKOIILYPLHGDNPQIWLPLF 576

RESULT 9
AAP50166
ID AAP50166 standard; Protein: 565 AA.
XX
AC AAP50166;
XX
DT 16-OCT-1991 (first entry)
DT
XX
DE Sequence of preprotricin encoded by prCL617.
XX
KM Toxin; anti-tumour therapy.
XX

```



XX The sequence relates to preprotricin protein encoded by the DNA sequence  
 CC given in A164137. The invention relates to a novel toxin (e.g., ricin)  
 CC based antiviral agent which is toxic to virus-infected cells, but  
 CC non-toxic to uninfected cells. The invention has anti-HIV and virocidal  
 CC activities. Its mechanism of action is through inactivation of cellular  
 CC ribosomes and enhancement of binding of the antiviral agent to galactose  
 CC residues on cell surfaces, and its cellular internalisation. The  
 CC invention is useful for treating human immunodeficiency virus infection  
 CC and other viral infections, especially retroviral infections. The  
 CC antiviral agent is activated in viral particles or early-stage infected  
 CC cells, killing the cells upon infection and effectively preventing the  
 CC integration of the viral genome into the host genome thereby preventing  
 CC the latency/rebound problem. The agent enters all HIV susceptible cells,  
 CC and not just cells known to act as host cells for the virus. The  
 CC antiviral agent remains inert in a cell until degraded in it, unless the  
 CC cell is infected with the virus, where the viral protease activates it.

XX Sequence 565 AA:

Query Match 97.7%; Score 2980; DB 22; Length 565;

Best Local Similarity 99.6%; Pred. No. 2,1e-246; Mismatches 1; Indels 0; Gaps 0;

Matches 563; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 MYAVATWLCFGSTGSMSTLEDNNIFPKOYPIINFTAGATVQSYTNFIRAVGRRLTGA 71  
 DB 1 MYAVATWLCFGSTGSMSTLEDNNIFPKOYPIINFTAGATVQSYTNFIRAVGRRLTGA 60  
 QY 72 DVHEHPIVLPNRYGLPIINORFLIVELSNHAEISVTALDVTNNAVYVGRAGNSAYFFHPD 131  
 DB 61 DVHNDIPVLPNRYGLPIINORFLIVELSNHAEISVTALDVTNNAVYVGRAGNSAYFFHPD 120  
 QY 132 NOEDAEATHTLFTDVONRYTFAGNGYDRLEQLAGNIRENIELNGPLEEASISALYYST 191  
 DB 121 NOEDAEATHTLFTDVONRYTFAGNGYDRLEQLAGNIRENIELNGPLEEASISALYYST 180  
 QY 192 GGTQPLTLARSFILCIOMISEARFOYIEGEMTRIRYRRSADPVSITLNSWGRLST 251  
 DB 181 GGTQPLTLARSFILCIOMISEARFOYIEGEMTRIRYRRSADPVSITLNSWGRLST 240  
 QY 252 AIOESNOCASFPIQLORRNGSKFSYDVSILPIIALMWYRCAPPSSQFSLLIRPVVP 311  
 DB 241 AIOESNOCASFPIQLORRNGSKFSYDVSILPIIALMWYRCAPPSSQFSLLIRPVVP 300  
 QY 312 NFNADVCMDEPIRIVIRIVGRNGLCVDRGRFHNGNAIQLMPCKSNTDANOQMTLRKNDTI 371  
 DB 301 NFNADVCMDEPIRIVIRIVGRNGLCVDRGRFHNGNAIQLMPCKSNTDANOQMTLRKNDTI 360  
 QY 372 RSNKGLCTTGYSPGYVYMIYDCNTAATDATRMOIWDNGTIINPRSSVLAAATSGNSGTT 431  
 DB 361 RSNKGLCTTGYSPGYVYMIYDCNTAATDATRMOIWDNGTIINPRSSVLAAATSGNSGTT 420  
 QY 432 LTVQNTIYAVSOGWLPNTNTOPTVTTIVGLYGLCLQANSOGVWIECCSSKAEQOMALYA 491  
 DB 421 LTVQNTIYAVSOGWLPNTNTOPTVTTIVGLYGLCLQANSOGVWIECCSSKAEQOMALYA 480  
 QY 492 DGSIRPOQRNDCLTSDSNIRETVYKILSCGPASSGQRMFKNDGTILNLVYGLVLDVRA 551  
 DB 481 DGSIRPOQRNDCLTSDSNIRETVYKILSCGPASSGQRMFKNDGTILNLVYGLVLDVRA 540  
 QY 552 SDPSLKQIITLPLHGPNOIWLPLF 576  
 DB 541 SDPSLKQIITLPLHGPNOIWLPLF 565

RESULT 11

AAP60240 standard; protein; 565 AA.

AC AAP60240:

XX 25-MAR-2003 (updated)  
 DT 02-JUL-1991 (first entry)

XX Preprotricin.  
 DE Preprotricin: antitumour agent;  
 XX Ricinus communis.  
 OS Ricinus communis.  
 PN EPI69006-A.  
 XX 22-JAN-1986.  
 PD 02-JUL-1985; 8SEP-0304711.  
 PF 13-JUL-1984; 84GB-0017915.  
 PR (ICIL ) IMPERIAL CHEM IND PLC.  
 PA Lord JM, Roberts LM, Atherton KT, Sharpe GS, Windass JD;  
 PI Tonge DW;  
 PI WPI: 1986-022682/04.  
 DR N-PSDB; AAN60193.  
 DR New biologically pure pro-ricin and (truncated) B-chain of ricin  
 PT - useful in mfg. ricin derivs. for more selective antitumour use  
 PT and obtd. by recombinant DNA procedures  
 PS Disclosure: Fig. 1; 31pp: English.

Preprotricin is formed in Ricinus communis in the biosynthesis of the  
 CC ricin toxin. Preprotricin is useful on modification giving a material  
 CC useful as an antitumour agent without the indiscriminate blinding of  
 CC ricin B chain.  
 CC (updated on 25-MAR-2003 to correct PA field.)

XX Sequence 565 AA:

Query Match 97.3%; Score 2968; DB 7; Length 565;

Best Local Similarity 99.3%; Pred. No. 2,3e-245; Mismatches 3; Indels 0; Gaps 0;

Matches 561; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 12 MYAVATWLCFGSTGSMSTLEDNNIFPKOYPIINFTAGATVQSYTNFIRAVGRRLTGA 71  
 DB 1 MYAVATWLCFGSTGSMSTLEDNNIFPKOYPIINFTAGATVQSYTNFIRAVGRRLTGA 60  
 QY 72 DVHEHPIVLPNRYGLPIINORFLIVELSNHAEISVTALDVTNNAVYVGRAGNSAYFFHPD 131  
 DB 61 DVHNDIPVLPNRYGLPIINORFLIVELSNHAEISVTALDVTNNAVYVGRAGNSAYFFHPD 120  
 QY 132 NOEDAEATHTLFTDVONRYTFAGNGYDRLEQLAGNIRENIELNGPLEEASISALYYST 191  
 DB 121 NOEDAEATHTLFTDVONRYTFAGNGYDRLEQLAGNIRENIELNGPLEEASISALYYST 180  
 QY 192 GGTQPLTLARSFILCIOMISEARFOYIEGEMTRIRYRRSADPVSITLNSWGRLST 251  
 DB 181 GGTQPLTLARSFILCIOMISEARFOYIEGEMTRIRYRRSADPVSITLNSWGRLST 240  
 QY 252 AIOESNOCASFPIQLORRNGSKFSYDVSILPIIALMWYRCAPPSSQFSLLIRPVVP 311  
 DB 241 AIOESNOCASFPIQLORRNGSKFSYDVSILPIIALMWYRCAPPSSQFSLLIRPVVP 300  
 QY 312 NFNADVCMDEPIRIVIRIVGRNGLCVDRGRFHNGNAIQLMPCKSNTDANOQMTLRKNDTI 371  
 DB 301 NFNADVCMDEPIRIVIRIVGRNGLCVDRGRFHNGNAIQLMPCKSNTDANOQMTLRKNDTI 360  
 QY 372 RSNKGLCTTGYSPGYVYMIYDCNTAATDATRMOIWDNGTIINPRSSVLAAATSGNSGTT 431  
 DB 361 RSNKGLCTTGYSPGYVYMIYDCNTAATDATRMOIWDNGTIINPRSSVLAAATSGNSGTT 420  
 QY 432 LTVQNTIYAVSOGWLPNTNTOPTVTTIVGLYGLCLQANSOGVWIECCSSKAEQOMALYA 491  
 DB 421 LTVQNTIYAVSOGWLPNTNTOPTVTTIVGLYGLCLQANSOGVWIECCSSKAEQOMALYA 480



QY 492 DGSIRPOONRDNCITSDSNIRETVYKILSCGPASSGGRMMFKNDGTLINLVSGLVDVRA 551  
 DB 481 DGSIRPOONRDNCITSDSNIRETVYKILSCGPASSGGRMMFKNDGTLINLVSGLVDVRA 540  
 QY 552 SDPSLKQIILYPLHGDNPQIWLPLF 576  
 DB 541 SDPSLKQIILYPLHGDNPQIWLPLF 565  
 RESULT 12  
 AAP90079  
 ID AAP90079 standard; protein; 562 AA.  
 AC AAP90079;  
 XX 25-MAR-2003 (updated)  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Ricin D.  
 XX  
 KM Ricin D; Ricinus communis; castor beans; Zanthariensis variety;  
 KM modified; lectin binding removed; reduced cell binding  
 XX  
 OS Ricinus communis (caster beans).  
 XX  
 PN WC8904839-A.  
 PD 01-JUN-1989.  
 XX  
 PF 23-NOV-1988; 88WO-US04238.  
 XX  
 PR 24-NOV-1987; 87US-0124735.  
 XX  
 PA (GEMV ) GENETICS INST INC.  
 XX  
 PI Brown EL, Jones S;  
 XX  
 DR WFI; 1989-178366/24.  
 DR N-PSDB; AAN90068.  
 XX  
 PT Modified ricin molecules and toxin conjugates  
 PT 1; - in which the lectin binding function of the B chain  
 PT 1; removed or diminished to reduce cell binding.  
 XX  
 PS Disclosure; fig 1; 51pp; English.  
 XX  
 CC Ricin D (see corresp. AAN90068) found on the EcoRI-HindIII fragment  
 CC of DNA from Ricinu communis, Zanthariensis variety. Patent  
 CC discloses many modifications of ricin in which the lectin binding  
 CC function of the B chain is diminished or removed, and conjugation  
 CC to toxins to eliminate cell binding.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 XX Sequence 562 AA:  
 SQ  
 Query Match 97.0%; Score 2960; DB 10; Length 562;  
 Best Local Similarity 97.6%; Pred. No. 1.1e-244;  
 Matches 562; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

DB 181 EAIISALYYSTGGTOLPLTANSFIICIMISEAARFOYIEGEMRTIRYNRRSAPDSVI 240  
 QY 241 TLNSWGRSLTAIQESNOGARASPIQOLORRNGSKFSYDVSIILPIITAMRYCARPPSS 300  
 DB 241 TLNSWGRSLTAIQESNOGARASPIQOLORRNGSKFSYDVSIILPIITAMRYCARPPSS 300  
 QY 301 QPSLIRPVVPENADVCMDPEPIYRIYGRNGLCYVDGDRFHNGNAIQLMPCKSNTDAN 360  
 DB 301 QPSLIRPVVPENADVCMDPEPIYRIYGRNGLCYVDGDRFHNGNAIQLMPCKSNTDAN 360  
 QY 361 QLMTLKRDNTIRSNCKLITTYGSPGVYMIYDCNFAATDAATRMQIMDNGTIIINRSGLV 420  
 DB 361 QLMTLKRDNTIRSNCKLITTYGSPGVYMIYDCNFAATDAATRMQIMDNGTIIINRSGLV 420  
 QY 421 LAATSGNSGTTITVQNTINIAVSQGLPTNNQPFYTTIVGLYGLCLQANSQVWIECCSS 480  
 DB 421 LAATSGNSGTTITVQNTINIAVSQGLPTNNQPFYTTIVGLYGLCLQANSQVWIECCSS 466  
 QY 481 EKAEQOMALYADGSIPOONRDNCITSDSNIRETVYKILSCGPASSGGRMMFKNDGTLIN 540  
 DB 467 EKAEQOMALYADGSIPOONRDNCITSDSNIRETVYKILSCGPASSGGRMMFKNDGTLIN 526  
 QY 541 LYSGLVDVRA SDPSLKQIILYPLHGDNPQIWLPLF 576  
 DB 527 LYSGLVDVRA SDPSLKQIILYPLHGDNPQIWLPLF 562  
 RESULT 13  
 AAG78304  
 ID AAG78304 standard; Protein; 565 AA.  
 XX  
 AC AAG78304;  
 XX  
 DT 27-NOV-2001 (first entry)  
 XX  
 DE Modified castor bean preprotricin (SEQ ID 10).  
 XX  
 KM Castor bean plant; preprotricin; ricin; A chain; B chain;  
 KM human immunodeficiency virus infection; HIV; toxin; antiviral agent;  
 KM retroviral infection; anti-HIV; virucide; viral protease.  
 XX  
 OS Chimeric - Ricinus communis  
 OS  
 OS Chimeric - Human immunodeficiency virus type 2.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /label= Signal\_peptide  
 FT 25..565  
 FT /label= Protricin  
 FT /note= "Protricin consists of the ricin A chain, a linker  
 FT peptide, and the ricin B chain. Protricin is  
 FT proteolytically cleaved between the A chain and  
 FT the linker to yield mature ricin"  
 FT Protein 25..291  
 FT /label= Ricin\_A\_chain  
 FT /note= "N-glycosidase"  
 FT Peptide 292..303  
 FT /label= Linker\_peptide  
 FT 296..297  
 FT /label= HIV\_protease\_cleavage\_site  
 FT Protein 304..565  
 FT /label= Ricin\_B\_chain  
 FT /note= "Galactose/N-acetyl galactosamine-binding lectin"  
 XX  
 PN W0200160393-A1.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 15-FEB-2001; 2001WO-US05282.  
 XX  
 PR 16-FEB-2000; 2000US-0182759.  
 XX



PA (BECH-) BECHTEL BWXT IDAHO LLC.  
 XX Keener WK, Ward TE;  
 PI  
 XX  
 DR WPI: 2001-581908/65.  
 DR N-PSDB; AA164145.  
 XX  
 PT Novel composition comprising toxin e.g., ricin based antiviral compound  
 PT useful for treating viral infections such as human immunodeficiency  
 PT virus infection.  
 XX  
 XX Example 1; Page 59-63; 66pp; English.  
 XX  
 CC The sequence relates to the amino acid sequence of a modified prepropricin  
 CC protein encoded by AA164145. The invention relates to a novel toxin  
 CC (e.g. ricin) based antiviral agent which is toxic to virus-infected  
 CC cells, but non-toxic to uninfected cells. The invention has anti-HIV and  
 CC virucide activities. The agent is able to enter all HIV susceptible  
 CC cells, and not just cells known to act as host cells for the virus. The  
 CC antiviral agent remains inert in a cell unless the cell is infected  
 CC with the HIV virus, where the viral protease activates it. Ricin's  
 CC mechanism of action is through inactivation of cellular ribosomes and  
 CC enhancement of binding of the antiviral agent to galactose residues on  
 CC cell surfaces, and its cellular internalisation. The invention is useful  
 CC for treating human immunodeficiency virus infection and other viral  
 CC infections, especially retroviral infections. The antiviral agent is  
 CC activated in viral particles or early-stage infected cells, killing the  
 CC cells upon infection and effectively preventing the integration of the  
 CC viral genome into the host genome thereby preventing the latency/rebound  
 CC problem.  
 XX  
 XX Sequence 565 AA:  
 SQ  
 Query Match 96.6%; Score 2947; DB 22; Length 565;  
 Best Local Similarity 98.8%; Pred. No. 1.4e-243;  
 Matches 558; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 QY 12 MYAVATWLCFGSTGNSFTLEDNNIFPKQPIINFTAGATVQSYTNFIRAVRGLTGA 71  
 DB 1 MYAVATWLCFGSTGNSFTLEDNNIFPKQPIINFTAGATVQSYTNFIRAVRGLTGA 60  
 QY 72 DYRHEIPVLPNRYGVPINORFLVELSNHAEISVTLADVTNAYVYVGRAGSAVFPHD 131  
 DB 61 DYRHEIPVLPNRYGVPINORFLVELSNHAEISVTLADVTNAYVYVGRAGSAVFPHD 120  
 QY 132 NOEDAEATHTLFTDVONRTYFAFGNGYDRLEQLAGLRNITLGNCPLEAISALYYST 191  
 DB 121 NOEDAEATHTLFTDVONRTYFAFGNGYDRLEQLAGLRNITLGNCPLEAISALYYST 180  
 QY 192 GGTQLEPTLARSEFIICOMISEARFOYLEGEMRTIRYKRSAAPPSTVTTLENSMGRLT 251  
 DB 181 GGTQLEPTLARSEFIICOMISEARFOYLEGEMRTIRYKRSAAPPSTVTTLENSMGRLT 240  
 QY 252 AIOESNOGAFASPIQLOQRNGSKFSYDVSIIPITAMVYCAPPPSQESLLIRPVV 311  
 DB 241 AIOESNOGAFASPIQLOQRNGSKFSYDVSIIPITAMVYCAPPPSQESLLIRPVV 300  
 QY 312 NFENADVCMPDEPIVIRIVGRNGLCVYRDOGRFHNGNAIOLPCKSNTDANQMTLTKRDMT 371  
 DB 301 NFENADVCMPDEPIVIRIVGRNGLCVYRDOGRFHNGNAIOLPCKSNTDANQMTLTKRDMT 360  
 QY 372 RSNKGCLTTYGYSPGVYVYIDCMTAATDATRMQIMDNGTITNPRSSVLVATSGNSGT 431  
 DB 361 RSNKGCLTTYGYSPGVYVYIDCMTAATDATRMQIMDNGTITNPRSSVLVATSGNSGT 420  
 QY 432 LTVQNTIYAVSGMPLTNTQPFYTVYGLYGLCLOANSQGVWIDCSEKAEQOWALYA 491  
 DB 421 LTVQNTIYAVSGMPLTNTQPFYTVYGLYGLCLOANSQGVWIDCSEKAEQOWALYA 480  
 QY 492 DGSIRPOQRNDCLTSDSNIRETVVKILSCGPASSGORMFANDGTITNLVSGVLVDRA 551  
 DB 481 DGSIRPOQRNDCLTSDSNIRETVVKILSCGPASSGORMFANDGTITNLVSGVLVDRA 540

QY 552 SDPSLKQITLYPLHNDPQNIWLPF 576  
 DB 541 SDPSLKQITLYPLHNDPQNIWLPF 565  
 RESULT 14  
 AAM25143  
 ID AAM25143 standard; Protein: 540 AA.  
 XX  
 AC AAM25143;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 02-DEC-1997 (first entry)  
 DE  
 XX Castor oil plant agglutinin inactive precursor.  
 XX  
 KW Maize; proRIP; ribosome inactivating protein; alpha; beta subunit;  
 KW internal linker; Barley translation inhibitor; Trichosanthin;  
 KW Ricin A-chain; Abrin-A A-chain; Saporin; SRT-1; Luffin A; MAP;  
 KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;  
 KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;  
 KW post-translational modification; cancer; neoplasia; HIV; AIDS;  
 KW human immunodeficiency virus; acquired immune deficiency syndrome.  
 XX  
 OS Ricinus communis.  
 XX  
 PN US5646026-A.  
 PD 08-JUL-1997.  
 XX  
 PF 07-JUN-1995; 95US-0485286.  
 XX  
 PR 09-DEC-1992; 92US-0987927.  
 PR 11-JUN-1990; 90US-0535636.  
 PR 26-JAN-1995; 95US-0378761.  
 PR 07-JUN-1995; 95US-0485286.  
 XX  
 PA (DOWC ) DOWELANCO.  
 PI Hey TD, Morgan AER, Walsh TA;  
 XX  
 DR WPI: 1997-362934/33.  
 XX  
 PT DNA encoding pro-ribosome inactivating proteins - inactive  
 PT precursors of ribosome inactivating proteins; can be expressed in  
 PT eukaryotic cells without causing cell death  
 PS  
 PS Claim 4; Column 121-124; 186pp; English.  
 XX  
 CC AAM25143 shows a castor oil plant agglutinin protein which was  
 CC engineered to contain a selectively removable internal peptide linker  
 CC sequence separating the alpha and beta units. When separated the two  
 CC units regain activity and are capable of inactivating eukaryotic  
 CC ribosomes and hence preventing protein production. Many different  
 CC ribosome inhibitory proteins (RIPs) may be produced with an internal  
 CC linker including maize RIP, Trichosanthin, Ricin A-chain, Abrin-A A-chain  
 CC and Saporin. The RIPs can be used in the construction of therapeutic  
 CC toxins targeted to specific cells such as tumour cells via the  
 CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.  
 CC A further use is in HIV therapy (see US4869903). There is interest  
 CC in expressing RIP recombinantly in host eukaryotic cells, because of  
 CC the capacity to provide correct post-translational processing. However,  
 CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting  
 CC in cell death. Since the inactive RIP proteins are not cytotoxic to  
 CC eukaryotic cells, they can be recombinantly expressed in such cells and  
 CC then converted to active RIP proteins.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC  
 XX  
 SQ Sequence 540 AA;  
 Query Match 83.0%; Score 2531.5; DB 18; Length 540;  
 Best Local Similarity 89.1%; Pred. No. 5.3e-208;  
 Matches 482; Conservative 26; Mismatches 32; Indels 1; Gaps 1;



Tue Sep 16 12:27:45 2003

us-10-083-336a-1.rag

Page 13

Db 540 F 540

Search completed: September 16, 2003, 11:45:15  
Job time : 98.9159 secs

...

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:42:03; Search time 32.8133 seconds  
(without alignments)  
742.718 Million cell updates/sec

Title: US-10-083-336A-1

Perfect score: 3051  
Sequence: 1 MKPGNTIVIMYAVATWLC.....KQILYPLHGPNOIWLPLF 576

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2531.5	83.0	540	1	US-08-378-761A-77
2	2531.5	83.0	540	1	US-08-485-286-77
3	1494.5	49.0	290	1	US-08-378-761A-27
4	1494.5	49.0	290	1	US-08-485-286-27
5	1494.5	49.0	290	6	5248606-4
6	1481.5	48.6	534	2	US-08-356-786-10
7	1372	45.0	267	1	US-07-901-707-1
8	1372	45.0	267	1	US-07-988-430-1
9	1372	45.0	267	1	US-08-425-336-1
10	1372	45.0	267	1	US-08-488-113B-1
11	1372	45.0	267	1	US-08-477-484B-1
12	1372	45.0	267	3	US-08-646-360-1
13	1372	45.0	267	3	US-08-839-765-1
14	1372	45.0	267	3	US-09-136-389-1
15	1372	45.0	267	4	US-09-610-838-1
16	1372	45.0	267	5	PCITUS-09487-1
17	1372	45.0	268	2	US-08-356-786-8
18	1362	44.6	267	1	US-08-218-303-16
19	1362	44.6	267	2	US-08-338-793D-61
20	1362	44.6	267	4	US-09-538-873-1
21	1295	42.4	263	3	US-08-776-059-35
22	906	29.7	263	3	US-08-776-059-43
23	906	29.7	264	3	US-08-776-059-33
24	428	14.0	282	1	US-08-324-301-15
25	423	13.9	251	4	US-09-538-873-3
26	421	13.8	267	1	US-08-378-761A-74
27	421	13.8	267	1	US-08-485-286-74

28	416	13.6	247	1	US-08-488-113B-6	Sequence 6, Appl
29	416	13.6	247	1	US-08-477-484B-6	Sequence 6, Appl
30	416	13.6	247	2	US-08-646-360-6	Sequence 6, Appl
31	416	13.6	247	3	US-08-839-765-6	Sequence 6, Appl
32	416	13.6	247	3	US-09-136-389-6	Sequence 6, Appl
33	416	13.6	247	4	US-09-610-838-6	Sequence 6, Appl
34	416	13.6	289	1	US-07-923-692C-4	Sequence 4, Appl
35	416	13.6	289	1	US-08-184-237-4	Sequence 4, Appl
36	416	13.6	289	2	US-08-482-920-4	Sequence 4, Appl
37	416	13.6	289	3	US-08-484-341-4	Sequence 4, Appl
38	416	13.6	289	3	US-08-483-502-4	Sequence 4, Appl
39	416	13.6	289	4	US-09-726-651A-4	Sequence 4, Appl
40	414.5	13.6	250	1	US-08-378-761A-71	Sequence 71, Appl
41	414.5	13.6	250	1	US-08-485-286-71	Sequence 71, Appl
42	401	13.1	496	3	US-08-902-486-12	Sequence 15, Appl
43	393	12.9	290	2	US-08-245-754A-2	Sequence 2, Appl
44	393	12.9	290	2	US-08-597-731-2	Sequence 2, Appl
45	389	12.7	248	3	US-08-902-486-7	Sequence 7, Appl

## ALIGNMENTS

RESULT 1  
US-08-378-761A-77  
Sequence 77, Application US/08378761A  
Patent No. 5635384  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/378,761A  
FILING DATE: 26-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 77:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 540 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-378-761A-77

Query Match 83.0%; Score 2531.5; DB 1; Length 540;  
Best Local Similarity 89.1%; Pred. No. 1.4e-238;  
Matches 482; Conservative 26; Mismatches 32; Indels 1; Gaps 1;

QY 36 IFPRQYIIINFTAGAVOSTYTFIRAVRRLTTGADVREIYVLPNRGCLPINOFTIV 95  
DB 1 IFPRQYPIINFTADATVESYTFNIRAVRSHLTTGADVREIYVLPNRGCLPISOFTIV 60

```

QY 96 ELSNHAELSVTLALDVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVQNRRTFAFG 155
DB 61 ELSNHAELSVTLALDVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVQNRRTFAFG 120
QY 156 GNDRLLEQLAGNIRENIELGNGPLEAISAALYYSTGGTQLPTLARSFTICQIMISEAR 215
DB 121 GNDRLLEQL-GGIRENIELGTGLEAISAALYYSTGGTQLPTLARSFTICQIMISEAR 179
QY 216 FOYIEGEMTRIRYRNRSAADPSVITLNSWGRSLTAIOESNOGAFASPIOLORRNSKF 275
DB 180 FOYIEGEMTRIRYRNRSAADPSVITLNSWGRSLTAIOESNOGAFASPIOLORRNSKF 239
QY 276 SYVDVSLITPITALMVRCAAPPPSSQFSLIRPVVFNPNADVCMDEPIVIRVGRNGLCV 335
DB 240 NYDVSILIPITALMVRCAAPPPSSQFSLIRPVVFNPNADVCMDEPIVIRVGRNGLCV 299
QY 336 DVRDGRFHNGNAIOLMPCKSNTDANQMTLKRDNTRNSGKCLTYYGSPGVYVMIYDCN 395
DB 300 DVTGEFFDGNPIOLMPCKSNTDANQMTLKRDNTRNSGKCLTYYGSPGVYVMIYDCN 359
QY 396 TAATDATRMOIMDNGTIIINPRSSVLAAISGSGTTLVQTNITAVSOGMLPTNNTQPEV 455
DB 360 TAATDATRMOIMDNGTIIINPRSSVLAAISGSGTTLVQTNITAVSOGMLPTNNTQPEV 419
QY 456 TTVVGLYGLCLQANSQGVWLEDCSEKAEQOMALYADGSIROPQNRDNCITSDSNIREFY 515
DB 420 TTVVGLYGLCLQANSQGVWLEDCSEKAEQOMALYADGSIROPQNRDNCITSDSNIREFY 479
QY 516 VKILSCGPASSGQRMFKNDGTLNLVYGLVDVRAADPSLKOIILVPLHGPNOIWLPL 575
DB 480 VKILSCGPASSGQRMFKNDGTLNLVYGLVDVRAADPSLKOIILVPLHGPNOIWLPL 539
QY 576 F 576
DB 540 F 540

RESULT 2
US-08-485-286-77
; Sequence 27, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651

```

```

; REFERENCE/DOCKET NUMBER: 382728
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-485-286-77

Query Match      83.0%; Score 2531.5; DB 1; Length 540;
Best Local Similarity 89.1%; Pred. No. 1.4e-238;
Matches 482; Conservative 26; Mismatches 32; Indels 1; Gaps 1;

QY 36 IFPKQYPIINFTTACATVOSTTNTFRVARGTLTGADVRHEIPLVLRVGLPIQNRRLV 95
DB 1 IFPKQYPIINFTTADATVESYTNFIRAVRSHLTGADVRIEPLVLRVGLPIQNRRLV 60
QY 96 ELSNHAELSVTLALDVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVQNRRTFAFG 155
DB 61 ELSNHAELSVTLALDVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVQNRRTFAFG 120
QY 156 GNDRLLEQLAGNIRENIELGNGPLEAISAALYYSTGGTQLPTLARSFTICQIMISEAR 215
DB 121 GNDRLLEQL-GGIRENIELGTGLEAISAALYYSTGGTQLPTLARSFTICQIMISEAR 179
QY 216 FOYIEGEMTRIRYRNRSAADPSVITLNSWGRSLTAIOESNOGAFASPIOLORRNSKF 275
DB 180 FOYIEGEMTRIRYRNRSAADPSVITLNSWGRSLTAIOESNOGAFASPIOLORRNSKF 239
QY 276 SYVDVSLITPITALMVRCAAPPPSSQFSLIRPVVFNPNADVCMDEPIVIRVGRNGLCV 335
DB 240 NYDVSILIPITALMVRCAAPPPSSQFSLIRPVVFNPNADVCMDEPIVIRVGRNGLCV 299
QY 336 DVRDGRFHNGNAIOLMPCKSNTDANQMTLKRDNTRNSGKCLTYYGSPGVYVMIYDCN 395
DB 300 DVTGEFFDGNPIOLMPCKSNTDANQMTLKRDNTRNSGKCLTYYGSPGVYVMIYDCN 359
QY 396 TAATDATRMOIMDNGTIIINPRSSVLAAISGSGTTLVQTNITAVSOGMLPTNNTQPEV 455
DB 360 TAATDATRMOIMDNGTIIINPRSSVLAAISGSGTTLVQTNITAVSOGMLPTNNTQPEV 419
QY 456 TTVVGLYGLCLQANSQGVWLEDCSEKAEQOMALYADGSIROPQNRDNCITSDSNIREFY 515
DB 420 TTVVGLYGLCLQANSQGVWLEDCSEKAEQOMALYADGSIROPQNRDNCITSDSNIREFY 479
QY 516 VKILSCGPASSGQRMFKNDGTLNLVYGLVDVRAADPSLKOIILVPLHGPNOIWLPL 575
DB 480 VKILSCGPASSGQRMFKNDGTLNLVYGLVDVRAADPSLKOIILVPLHGPNOIWLPL 539
QY 576 F 576
DB 540 F 540

RESULT 3
US-08-378-761A-27
; Sequence 27, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS

```

STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/378,761A  
FILING DATE: 26-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 290 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-378-761A-27

Query Match 49.0%; Score 1494.5; DB 1; Length 290;  
Best Local Similarity 99.7%; Pred. No. 1.le-137;  
Matches 290; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 12 MYAAATWLCFGSTGSGMSTFLEDDNNIFPKOYPIINFTTAGATVOSYTNFIRAVRGLTTGA 71  
DB 1 MYAAATWLCFGSTGSGMSTFLEDDNNIFPKOYPIINFTTAGATVOSYTNFIRAVRGLTTGA 60  
QY 72 DVREHPIVLPNRYGLPINORFLLVLSNHAELSVTLADVTNAYVYGRAGNSAYFFHPD 131  
DB 61 DVREHPIVLPNRYGLPINORFLLVLSNHAELSVTLADVTNAYVYGRAGNSAYFFHPD 120  
QY 132 NOEDAEATHTLFTDVONRYTFAGGNYDRLEQLAGNLRENIELGNGPLEEASISALYYST 191  
DB 121 NOEDAEATHTLFTDVONRYTFAGGNYDRLEQLAGNLRENIELGNGPLEEASISALYYST 180  
QY 192 GGTOLPTLARSFIICIMISEARFOYIEGEMRTIRYNRRSADPSVITLNSMGRIST 251  
DB 181 GGTOLPTLARSFIICIMISEARFOYIEGEMRTIRYNRRSADPSVITLNSMGRIST 240  
QY 252 AIOESNOGAFASPIOLORRNGSKFSYDVSVILIPITIALMYRCAPPSSQF 302  
DB 241 AIOESNOGAFASPIOLORRNGSKFSYDVSVILIPITIALMYRCAPP-SQF 290

RESULT 4  
US-08-485-286-27  
Sequence 27, Application US/08485286  
Patent No. 5646026  
Patent No. 5646026 5646119  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,286  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/378761  
FILING DATE: 26-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 290 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-286-27

Query Match 49.0%; Score 1494.5; DB 1; Length 290;  
Best Local Similarity 99.7%; Pred. No. 1.le-137;  
Matches 290; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 12 MYAAATWLCFGSTGSGMSTFLEDDNNIFPKOYPIINFTTAGATVOSYTNFIRAVRGLTTGA 71  
DB 1 MYAAATWLCFGSTGSGMSTFLEDDNNIFPKOYPIINFTTAGATVOSYTNFIRAVRGLTTGA 60  
QY 72 DVREHPIVLPNRYGLPINORFLLVLSNHAELSVTLADVTNAYVYGRAGNSAYFFHPD 131  
DB 61 DVREHPIVLPNRYGLPINORFLLVLSNHAELSVTLADVTNAYVYGRAGNSAYFFHPD 120  
QY 132 NOEDAEATHTLFTDVONRYTFAGGNYDRLEQLAGNLRENIELGNGPLEEASISALYYST 191  
DB 121 NOEDAEATHTLFTDVONRYTFAGGNYDRLEQLAGNLRENIELGNGPLEEASISALYYST 180  
QY 192 GGTOLPTLARSFIICIMISEARFOYIEGEMRTIRYNRRSADPSVITLNSMGRIST 251  
DB 181 GGTOLPTLARSFIICIMISEARFOYIEGEMRTIRYNRRSADPSVITLNSMGRIST 240  
QY 252 AIOESNOGAFASPIOLORRNGSKFSYDVSVILIPITIALMYRCAPPSSQF 302  
DB 241 AIOESNOGAFASPIOLORRNGSKFSYDVSVILIPITIALMYRCAPP-SQF 290

RESULT 5  
5248606-4  
Patent No. 5248606  
APPLICANT: WALSH, TERENCE A.; HEY, TIMOTHY D.; MORGAN,  
ALICE E.R.  
TITLE OF INVENTION: DNA ENCODING INACTIVE PRECURSOR AND  
ACTIVE FORMS OF MAIZE RIBOSOME INACTIVATIN  
NUMBER OF SEQUENCES: 49  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/535,636  
FILING DATE: 11-JUN-1990  
SEQ ID NO:4:  
LENGTH: 290  
5248606-4

Query Match 49.0%; Score 1494.5; DB 6; Length 290;  
Best Local Similarity 99.7%; Pred. No. 1.le-137;  
Matches 290; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 12 MYAAATWLCFGSTGSGMSTFLEDDNNIFPKOYPIINFTTAGATVOSYTNFIRAVRGLTTGA 71  
DB 1 MYAAATWLCFGSTGSGMSTFLEDDNNIFPKOYPIINFTTAGATVOSYTNFIRAVRGLTTGA 60

QY 72 DVHREIPLRVKGLPTINORFIIVELSNHAEIYTLALDVTNAVYVGRAGNSAYFFHPD 131  
DB 61 DVHREIPLRVKGLPTINORFIIVELSNHAEIYTLALDVTNAVYVGRAGNSAYFFHPD 120  
QY 132 NOEAEATITHEFTVQNRRTFAFGANDRLEOLAGNIRENIEGNGLEAISALYYST 191  
DB 121 NOEAEATITHEFTVQNRRTFAFGANDRLEOLAGNIRENIEGNGLEAISALYYST 180  
QY 192 GGTOPLTARSFTICIMISEARFOYIEGEMTRIRYNRSAPDPSVITLENSWGRUST 251  
DB 181 GGTOPLTARSFTICIMISEARFOYIEGEMTRIRYNRSAPDPSVITLENSWGRUST 240  
QY 252 AIOESNOGAFSPITOLORRNGSKFSYDVSTILPIITLAMYRCAPPSSQF 302  
DB 241 AIOESNOGAFSPITOLORRNGSKFSYDVSTILPIITLAMYRCAPPSSQF 290

## RESULT 6

US-08-356-786-10

Sequence 10, Application US/08356786

Patent No. 5877305

GENERAL INFORMATION:

APPLICANT: Huston, James S.

APPLICANT: Oppermann, Hermann

APPLICANT: Houston, L. L.

APPLICANT: Ring, David B.

TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESS: Edmund R. Pitcher, Testa, Hurwitz, &amp; Thibault

STREET: Exchange Place, 53 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/356,786

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/831,967

FILING DATE: 06-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: CRP-053

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 534 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-356-786-10

Query Match 48.6%; Score 1481.5; DB 2; Length 534;

Best Local Similarity 72.9%; Pred. No. 5.8e-136;

Matches 312; Conservative 26; Mismatches 41; Indels 49; Gaps 11;

QY 36 IFPKOYPIINTTGTAGATVOSTNFTIRAVRGTLTGADVRHEIPLRVKGLPTINORFIIV 95  
DB 4 IFPKOYPIINTTGTAGATVOSTNFTIRAVRGTLTGADVRHEIPLRVKGLPTINORFIIV 63  
QY 96 ELSNHAEIYTLALDVTNAVYVGRAGNSAYFFHPDNOEAEATITHEFTVQNRRTFAFG 155

DB 64 ELSNHAEIYTLALDVTNAVYVGRAGNSAYFFHPDNOEAEATITHEFTVQNRRTFAFG 123  
QY 156 GNVDRLEOLAGNIRENIEGNGLEAISALYYSTGGTQPLTARSFTICIMISEAR 215  
DB 124 GNVDRLEOLAGNIRENIEGNGLEAISALYYSTGGTQPLTARSFTICIMISEAR 183  
QY 216 FOYIEGEMTRIRYNRSAPDPSVITLENSWGRUSTAIOESNOGAFSPITOLORRNGSKF 275  
DB 184 FOYIEGEMTRIRYNRSAPDPSVITLENSWGRUSTAIOESNOGAFSPITOLORRNGSKF 243  
QY 276 SYVDVSIILPIITLAMYRCAPPSSQFSLIRPVVFNENADVCMPDEPVRIVGRNGLCV 335  
DB 244 SYVDVSIILPIITLAMYRCAPPSSQFSLIRPVVFNENADVCMPDEPVRIVGRNGLCV 298  
QY 336 DVROGFRHNGNAIQMPCKSN--TDANOLWTLKRDNTIRSNKCLTYYGY-----SPGVY 388  
DB 299 ELK---KPGETVKI--SCKASGYTFANQGMNMMK---QADPGKLMKMGWINTYTGOSTY 349  
QY 389 V-----MYDCNTATDTRMGI-----WDNGTIINPSSSLVL 421  
DB 350 ADDPKERFAFSLETSAITTA-HIQINNLRNEDSATYFCARREGFAIWGGITLVYSASISS 408  
QY 422 AATSGNSG 429  
DB 409 SGGGSGSG 416

## RESULT 7

US-07-901-707-1

Sequence 1, Application US/07901707

Patent No. 5376546

GENERAL INFORMATION:

APPLICANT: Bernhard, Susan L.

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Steve F.

APPLICANT: Lane, Julie A.

TITLE OF INVENTION: Materials Comprising and Methods of

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray &amp;

ADDRESS: Bicknell

STREET: Two First National Plaza, 20 South Clark

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/901,707

FILING DATE: 19920619

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: No. 5376546and, Greta E.

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 27129/30910

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 346-5750

TELEFAX: (312) 984-5750

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: AMINO ACID



TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-901-707-1

Query Match 45.0%; Score 1372; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 9,4e-126;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFFKQYPIINFTTACATVOSTNFIKAVRGRLTGADVHHEIPVLPNRYGLPINORFLV 95  
DB 1 IFFKQYPIINFTTACATVOSTNFIKAVRGRLTGADVHHEIPVLPNRYGLPINORFLV 60  
QY 96 ELSNHAELSVTLALDVTNAYVYGRAGNSAYFFHPDNOEDAELTHLFTDVQNNRYTFAFG 155  
DB 61 ELSNHAELSVTLALDVTNAYVYGRAGNSAYFFHPDNOEDAELTHLFTDVQNNRYTFAFG 120  
QY 156 GNYDRLEQLAGNLRENIEELGNGLPELAEISALYYSTGCTQLPTLARSFTICIMISEAR 215  
DB 121 GNYDRLEQLAGNLRENIEELGNGLPELAEISALYYSTGCTQLPTLARSFTICIMISEAR 180  
QY 216 FOYIEGEMRTIRYRNRSAAPDSVTITLENSWGRLSTAIOESNOGAFASPIQLORRNGSKF 275  
DB 181 FOYIEGEMRTIRYRNRSAAPDSVTITLENSWGRLSTAIOESNOGAFASPIQLORRNGSKF 240  
QY 276 SVYDVSLIPITIALMWYRCAPPSSQF 302  
DB 241 SVYDVSLIPITIALMWYRCAPPSSQF 267

RESULT 8  
US-07-988-430-1  
Sequence 1, Application us/07988430  
Patent No. 5416202

## GENERAL INFORMATION:

APPLICANT: Bernhardt, Susan L.  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Lane, Julie A.  
APPLICANT: Lei, Shau-Ping

TITLE OF INVENTION: Materials Comprising and Methods of  
Preparation and Use for Ribosome-Inactivating Proteins

NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell

STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/988,430  
FILING DATE: 19921209  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:

NAME: No. 5416202and, Greta E.  
REGISTRATION NUMBER: 35302  
REFERENCE/DOCKET NUMBER: 31133  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740

TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-988-430-1

Query Match 45.0%; Score 1372; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 9,4e-126;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFFKQYPIINFTTACATVOSTNFIKAVRGRLTGADVHHEIPVLPNRYGLPINORFLV 95  
DB 1 IFFKQYPIINFTTACATVOSTNFIKAVRGRLTGADVHHEIPVLPNRYGLPINORFLV 60  
QY 96 ELSNHAELSVTLALDVTNAYVYGRAGNSAYFFHPDNOEDAELTHLFTDVQNNRYTFAFG 155  
DB 61 ELSNHAELSVTLALDVTNAYVYGRAGNSAYFFHPDNOEDAELTHLFTDVQNNRYTFAFG 120  
QY 156 GNYDRLEQLAGNLRENIEELGNGLPELAEISALYYSTGCTQLPTLARSFTICIMISEAR 215  
DB 121 GNYDRLEQLAGNLRENIEELGNGLPELAEISALYYSTGCTQLPTLARSFTICIMISEAR 180  
QY 216 FOYIEGEMRTIRYRNRSAAPDSVTITLENSWGRLSTAIOESNOGAFASPIQLORRNGSKF 275  
DB 181 FOYIEGEMRTIRYRNRSAAPDSVTITLENSWGRLSTAIOESNOGAFASPIQLORRNGSKF 240  
QY 276 SVYDVSLIPITIALMWYRCAPPSSQF 302  
DB 241 SVYDVSLIPITIALMWYRCAPPSSQF 267

RESULT 9  
US-08-425-336-1  
Sequence 1, Application us/08425336  
Patent No. 5621083

## GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studilka, Gary M.  
APPLICANT: Lane, Julie A.  
APPLICANT: Lei, Shau-Ping

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins

NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
ADDRESSEE: Bicknell

STREET: 6300 Sears Tower, 233 South Wacker Drive  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/425,336  
FILING DATE: 18-APR-1995  
CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/901,707

APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Thomas C.  
REGISTRATION NUMBER: P-36,989  
REFERENCE/DOCKET NUMBER: 31394

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300  
 TELEFAX: 312/474-0448  
 TELEX: 25-3856  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 267 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-425-336-1

Query Match 45.0%; Score 1372; DB 1; Length 267;  
 Best local Similarity 100.0%; Pred. No. 9,4e-126;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINORFTLV 95  
 |||||||  
 DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINORFTLV 60  
 QY 96 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAG 155  
 |||||||  
 DB 61 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAG 120  
 QY 156 GNYDRLEQLAGNIRENIEIGNGPLEEASALYYSTGGTOLPTLARSFTICIMISEAR 215  
 |||||||  
 DB 121 GNYDRLEQLAGNIRENIEIGNGPLEEASALYYSTGGTOLPTLARSFTICIMISEAR 180  
 QY 216 FOYIEGEMRTIRYRRSAPDPSVITLNSWGRSLTAIOESNOGAFASPIOLORRNGSKF 275  
 |||||||  
 DB 181 FOYIEGEMRTIRYRRSAPDPSVITLNSWGRSLTAIOESNOGAFASPIOLORRNGSKF 240  
 QY 276 SVYDVSILPIITLIMVYRCAPPSSQF 302  
 |||||||  
 DB 241 SVYDVSILPIITLIMVYRCAPPSSQF 267

## RESULT 10

US-08-488-113B-1  
 ; Sequence 1, Application US/08488113B  
 ; Patent No. 5744580

## GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
 APPLICANT: Carroll, Stephen F.  
 APPLICANT: Studulka, Gary M.  
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
 TITLE OF INVENTION: Proteins  
 NUMBER OF SEQUENCES: 169  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: McAndrews, Heid & Malloy, Ltd.  
 STREET: 500 West Madison Street, 34th floor  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60661

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/488,113B  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/425,336  
 FILING DATE: 18-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/064,691  
 FILING DATE: 12-MAY-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/988,430  
 FILING DATE: 09-DEC-1992

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707  
 FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/787,567  
 FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:  
 NAME: McNicholas, Janet M.  
 REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 110220507/200-70.P3.C2A  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889  
 TELEFAX: 312/707-9155  
 TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear

MOLECULE TYPE: protein  
 US-08-488-113B-1

Query Match 45.0%; Score 1372; DB 1; Length 267;  
 Best local Similarity 100.0%; Pred. No. 9,4e-126;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINORFTLV 95  
 |||||||  
 DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINORFTLV 60  
 QY 96 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAG 155  
 |||||||  
 DB 61 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAG 120  
 QY 156 GNYDRLEQLAGNIRENIEIGNGPLEEASALYYSTGGTOLPTLARSFTICIMISEAR 215  
 |||||||  
 DB 121 GNYDRLEQLAGNIRENIEIGNGPLEEASALYYSTGGTOLPTLARSFTICIMISEAR 180  
 QY 216 FOYIEGEMRTIRYRRSAPDPSVITLNSWGRSLTAIOESNOGAFASPIOLORRNGSKF 275  
 |||||||  
 DB 181 FOYIEGEMRTIRYRRSAPDPSVITLNSWGRSLTAIOESNOGAFASPIOLORRNGSKF 240  
 QY 276 SVYDVSILPIITLIMVYRCAPPSSQF 302  
 |||||||  
 DB 241 SVYDVSILPIITLIMVYRCAPPSSQF 267

RESULT 11  
 US-08-477-484B-1  
 ; Sequence 1, Application US/08477484B  
 ; Patent No. 5756699

## GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
 APPLICANT: Carroll, Stephen F.  
 APPLICANT: Studulka, Gary M.  
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
 TITLE OF INVENTION: Proteins  
 NUMBER OF SEQUENCES: 169  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: McAndrews, Heid & Malloy, Ltd.  
 STREET: 500 West Madison Street, 34th floor  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60661

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/477,484B  
 FILING DATE: 07-JUN-1995

CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-477-484B-1

Query Match 45.0%; Score 1372; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 9,4e-126;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKOYPIINFTTAGATVQSYTNFIRAVRGRLLTGADVRHEIPVLPNRVGLPINORITLY 95  
DB 1 IFPKOYPIINFTTAGATVQSYTNFIRAVRGRLLTGADVRHEIPVLPNRVGLPINORITLY 60  
QY 96 ELSNNAELSVTLALDVTNAVYVGRAGNSAYFFHFDNQDEDAITHLFTDVQNRRTFAFG 155  
DB 61 ELSNNAELSVTLALDVTNAVYVGRAGNSAYFFHFDNQDEDAITHLFTDVQNRRTFAFG 120  
QY 156 GNYDRLEOLAGNLRNIEELGNGPLEAISALYYSTGGTQPLTARSFFICIONISEAR 215  
DB 121 GNYDRLEOLAGNLRNIEELGNGPLEAISALYYSTGGTQPLTARSFFICIONISEAR 180  
QY 216 FOYIEGEMRTIRRYNRSAPDPSVITLNSMGRSLTAIOESNOGAFASPIQLORRNSKF 275  
DB 181 FOYIEGEMRTIRRYNRSAPDPSVITLNSMGRSLTAIOESNOGAFASPIQLORRNSKF 240  
QY 276 SVYDVSILIPITIALMRYRCAPPSSQF 302  
DB 241 SVYDVSILIPITIALMRYRCAPPSSQF 267

## RESULT 12

US-08-646-360-1  
Sequence 1, Application US/08646360  
Patent No. 5837491  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
PROTEINS  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois

COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-360-1

Query Match 45.0%; Score 1372; DB 2; Length 267;  
Best Local Similarity 100.0%; Pred. No. 9,4e-126;  
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKOYPIINFTTAGATVQSYTNFIRAVRGRLLTGADVRHEIPVLPNRVGLPINORITLY 95  
DB 1 IFPKOYPIINFTTAGATVQSYTNFIRAVRGRLLTGADVRHEIPVLPNRVGLPINORITLY 60  
QY 96 ELSNNAELSVTLALDVTNAVYVGRAGNSAYFFHFDNQDEDAITHLFTDVQNRRTFAFG 155  
DB 61 ELSNNAELSVTLALDVTNAVYVGRAGNSAYFFHFDNQDEDAITHLFTDVQNRRTFAFG 120  
QY 156 GNYDRLEOLAGNLRNIEELGNGPLEAISALYYSTGGTQPLTARSFFICIONISEAR 215  
DB 121 GNYDRLEOLAGNLRNIEELGNGPLEAISALYYSTGGTQPLTARSFFICIONISEAR 180  
QY 216 FOYIEGEMRTIRRYNRSAPDPSVITLNSMGRSLTAIOESNOGAFASPIQLORRNSKF 275  
DB 181 FOYIEGEMRTIRRYNRSAPDPSVITLNSMGRSLTAIOESNOGAFASPIQLORRNSKF 240  
QY 276 SVYDVSILIPITIALMRYRCAPPSSQF 302  
DB 241 SVYDVSILIPITIALMRYRCAPPSSQF 267

## RESULT 13

US-08-839-765-1  
Sequence 1, Application US/08839765  
Patent No. 6146631  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.

```

APPLICANT: Carroll, Stephen F.
APPLICANT: Studinka, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
TITLE OF INVENTION: Proteins
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESS: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,765
FILING DATE: 15-APR-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/425,336
FILING DATE: 18-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-839-765-1

Query Match          45.0%; Score 1372; DB 3; Length 267;
Best Local Similarity 100.0%; Pred. No. 9,4e-126;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB          241 SVYDSILPIITLAWYRCAPPPSSQF 267
|||||
RESULT 14
US-09-136-389-1
Sequence 1, Application US/09136389
Patent No. 6146850
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Stephen F.
APPLICANT: Studinka, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
TITLE OF INVENTION: Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESS: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-136-389-1

Query Match          45.0%; Score 1372; DB 3; Length 267;
Best Local Similarity 100.0%; Pred. No. 9,4e-126;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db      61  ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNGDEAITHLFTDVQNNRYTFAFG 120
QY      156  GNYDRLEQLAGNLRNIEELGNGPLEEASALYYSTGTQPLTARSFITCIOMISEAR 215
Db      121  GNYDRLEQLAGNLRNIEELGNGPLEEASALYYSTGTQPLTARSFITCIOMISEAR 180
QY      216  FOYIEGEMRTIRYRNRSAAPDPSVITLNSWGRLSTAIOESNOGAFASPIQLORRNGSKF 275
Db      181  FOYIEGEMRTIRYRNRSAAPDPSVITLNSWGRLSTAIOESNOGAFASPIQLORRNGSKF 240
QY      276  SVYDVSLIPITIALMRYRCAPPPSSQF 302
Db      241  SVYDVSLIPITIALMRYRCAPPPSSQF 267

RESULT 15
US-09-610-838-1
; Sequence 1, Application US/09610838
; Patent No. 6376217
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnicka, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,838
; FILING DATE: 06-JUL-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/136,389
; FILING DATE: 18-AUG-1998
; APPLICATION NUMBER: 08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids

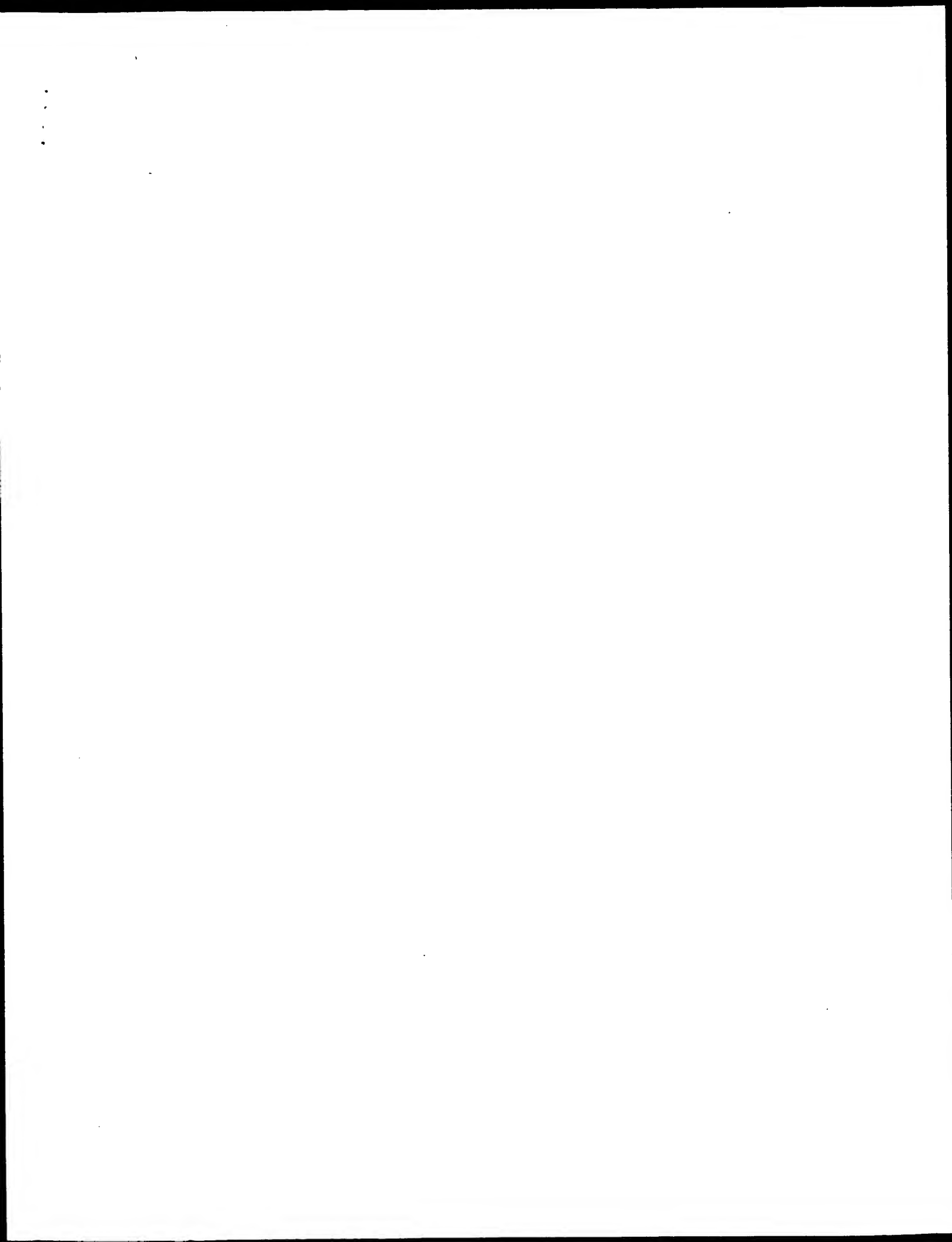
```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-610-838-1
Query Match      45.0%; Score 1372; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 9,4e-126;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      36  IFFKQYPIINFTTAGATVQSYTNFIRAVGRLLTGGADVHREIPLVLPNRYGLPINORFIIY 95
Db      1  IFFKQYPIINFTTAGATVQSYTNFIRAVGRLLTGGADVHREIPLVLPNRYGLPINORFIIY 60
QY      96  ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNGDEAITHLFTDVQNNRYTFAFG 155
Db      61  ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNGDEAITHLFTDVQNNRYTFAFG 120
QY      156  GNYDRLEQLAGNLRNIEELGNGPLEEASALYYSTGTQPLTARSFITCIOMISEAR 215
Db      121  GNYDRLEQLAGNLRNIEELGNGPLEEASALYYSTGTQPLTARSFITCIOMISEAR 180
QY      216  FOYIEGEMRTIRYRNRSAAPDPSVITLNSWGRLSTAIOESNOGAFASPIQLORRNGSKF 275
Db      181  FOYIEGEMRTIRYRNRSAAPDPSVITLNSWGRLSTAIOESNOGAFASPIQLORRNGSKF 240
QY      276  SVYDVSLIPITIALMRYRCAPPPSSQF 302
Db      241  SVYDVSLIPITIALMRYRCAPPPSSQF 267

```

Search completed: September 16, 2003, 11:51:45  
Job time : 34.8133 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:49:23 ; Search time 59.5688 Seconds  
(without alignments)  
1439.723 Million cell updates/sec

Title: US-10-083-336a-1  
Perfect score: 3051  
Sequence: 1 MKPGNTIVIMVAVATWLC.....KQILLYPLHGDPNOIMPLPLF 576

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 556269 seqs, 148893369 residues  
Total number of hits satisfying chosen parameters: 556269

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PC1\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PC1US\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10C\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1372	45.0	267	12	US-10-127-890-1
2	1362	44.6	267	12	US-10-282-935-1
3	906	29.7	263	9	US-09-347-064-10
4	906	29.7	267	9	US-09-347-064-4
5	423	13.9	251	12	US-10-282-935-3
6	421	13.8	247	10	US-09-792-793A-39
7	416	13.6	247	12	US-10-127-890-6
8	416	13.6	289	10	US-10-280-679B-4
9	389	12.7	247	10	US-09-792-793A-34
10	383	12.6	252	9	US-09-347-064-8
11	379	12.4	252	9	US-09-347-064-2
12	358	11.7	251	12	US-10-127-890-111
13	354	11.6	251	12	US-10-127-890-102
14	353.5	11.6	263	12	US-10-127-890-7
15	348.5	11.4	332	9	US-09-765-527-251

16	348	11.4	263	12	US-10-127-890-4	Sequence 4, Appl1
17	347	11.4	251	12	US-10-127-890-110	Sequence 110, App
18	347	11.4	293	9	US-09-765-527-259	Sequence 259, App
19	346	11.3	251	12	US-10-127-890-101	Sequence 101, App
20	345	11.3	251	12	US-10-127-890-99	Sequence 99, Appl1
21	345	11.3	251	12	US-10-127-890-104	Sequence 104, App
22	345	11.3	251	12	US-10-127-890-107	Sequence 107, App
23	345	11.3	309	9	US-09-765-527-253	Sequence 253, App
24	344.5	11.3	251	12	US-10-127-890-103	Sequence 103, App
25	344	11.3	251	12	US-10-127-890-100	Sequence 100, App
26	344	11.3	251	12	US-10-127-890-105	Sequence 105, App
27	344	11.3	251	12	US-10-127-890-106	Sequence 106, App
28	343	11.2	251	9	US-09-765-527-247	Sequence 247, App
29	343	11.2	251	12	US-10-127-890-2	Sequence 2, Appl1
30	341	11.2	251	12	US-10-127-890-108	Sequence 108, App
31	336	11.0	251	12	US-10-127-890-109	Sequence 109, App
32	321	10.5	248	12	US-10-127-890-5	Sequence 5, Appl1
33	305.5	10.0	314	10	US-09-978-274A-2	Sequence 2, Appl1
34	295.5	9.7	261	12	US-10-127-890-9	Sequence 9, Appl1
35	291	9.5	263	10	US-09-978-274A-4	Sequence 4, Appl1
36	246	8.1	44	15	US-10-137-077-17	Sequence 17, Appl1
37	216	7.1	41	15	US-10-137-077-18	Sequence 18, Appl1
38	216	7.1	250	10	US-09-792-793A-36	Sequence 36, Appl1
39	216	7.1	250	10	US-10-127-890-8	Sequence 8, Appl1
40	214	7.0	145	14	US-10-074-527-5	Sequence 5, Appl1
41	192	6.3	110	10	US-09-978-274A-8	Sequence 8, Appl1
42	184	6.0	420	15	US-10-156-761-14970	Sequence 14970, A
43	167.5	5.5	647	15	US-10-156-761-10246	Sequence 10246, A
44	159	5.2	658	15	US-10-156-761-9724	Sequence 9724, Ap
45	157.5	5.2	492	9	US-09-770-621-4	Sequence 4, Appl1

#### ALIGNMENTS

RESULT 1  
US-10-127-890-1  
; Sequence 1, Application US/10127890  
; Publication No. US20030166196A1  
GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
Stadl, Stephen F.

Studinka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor

CITY: Chicago  
STATE: Illinois  
COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996

APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994

APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787,567  
 FILING DATE: 04-NOV-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: McNicholas, Janet M.  
 REGISTRATION NUMBER: 32,918  
 REFERENCE/DOCKET NUMBER: 200-70,704  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312/707-8889  
 TELEFAX: 312/707-9155  
 TELEEX: 650 388-1248  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 267 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-10-127-890-1

Query Match 45.0%; Score 1372; DB 12; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 3,1e-126;  
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTTAGATVQSYTFNFIKAVRGRLFTGADVREHIEIPVLPNVRGLPINQRFILV 95  
 DB 1 IFPKQYPIINFTTAGATVQSYTFNFIKAVRGRLFTGADVREHIEIPVLPNVRGLPINQRFILV 60  
 QY 96 ELSNHAELSVTLADVTNAVAVGYRAGNSAYFFHPDQEDAEATITLFTDVONRTFFAGN 155  
 DB 61 ELSNHAELSVTLADVTNAVAVGYRAGNSAYFFHPDQEDAEATITLFTDVONRTFFAGN 120  
 QY 156 GNYRLBOLAGNLENIEINGPLEAISALYYSTGTLPTLARSFTICLQIMISEAAR 215  
 DB 121 GNYRLBOLAGNLENIEINGPLEAISALYYSTGTLPTLARSFTICLQIMISEAAR 180  
 QY 216 FQYIEGMRTRIRYRRSADPDSYITLNSWGRSLTAIOESNOGAFASPIQLORRNGSKF 275  
 DB 181 FQYIEGMRTRIRYRRSADPDSYITLNSWGRSLTAIOESNOGAFASPIQLORRNGSKF 240  
 QY 276 SVYDVSILPIIALMVRCAAPPSQF 302  
 DB 241 SVYDVSILPIIALMVRCAAPPSQF 267

# RESULT 2 US-10-282-935-1

Sequence 1, Application US/10282935  
 Publication No. US20030143193A1

GENERAL INFORMATION:  
 APPLICANT: VITETTA, ELLEN S.  
 APPLICANT: GHETIE, VICTOR F.  
 APPLICANT: SMALLSHAW, JOAN  
 APPLICANT: BALUNA, ROXANA G.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF  
 TITLE OF INVENTION: PROTEINACEOUS COMPOUNDS  
 FILE REFERENCE: URS: 88405  
 CURRENT APPLICATION NUMBER: US/10/282,935  
 CURRENT FILING DATE: 2002-10-29  
 PRIOR APPLICATION NUMBER: 09/538,873  
 PRIOR FILING DATE: 2000-03-30  
 PRIOR APPLICATION NUMBER: 60/126,826  
 PRIOR FILING DATE: 1999-03-30  
 NUMBER OF SEQ ID NOS: 23  
 SOFTWARE: Patentln Ver. 2.1  
 SEQ ID NO: 1

LENGTH: 267  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 OTHER INFORMATION: Peptide  
 US-10-282-935-1

Query Match 44.6%; Score 1362; DB 12; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 3e-125;  
 Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 PKQYPIINFTTAGATVQSYTFNFIKAVRGRLFTGADVREHIEIPVLPNVRGLPINQRFILV 97  
 DB 3 PKQYPIINFTTAGATVQSYTFNFIKAVRGRLFTGADVREHIEIPVLPNVRGLPINQRFILV 62  
 QY 98 SNHAELSVTLADVTNAVAVGYRAGNSAYFFHPDQEDAEATITLFTDVONRTFFAGN 157  
 DB 63 SNHAELSVTLADVTNAVAVGYRAGNSAYFFHPDQEDAEATITLFTDVONRTFFAGN 122  
 QY 156 YRLBOLAGNLENIEINGPLEAISALYYSTGTLPTLARSFTICLQIMISEAARFQ 217  
 DB 123 YRLBOLAGNLENIEINGPLEAISALYYSTGTLPTLARSFTICLQIMISEAARFQ 182  
 QY 218 YIEGMRTRIRYRRSADPDSYITLNSWGRSLTAIOESNOGAFASPIQLORRNGSKFV 277  
 DB 183 YIEGMRTRIRYRRSADPDSYITLNSWGRSLTAIOESNOGAFASPIQLORRNGSKFV 242  
 QY 278 YDVSILPIIALMVRCAAPPSQF 302  
 DB 243 YDVSILPIIALMVRCAAPPSQF 267

# RESULT 3 US-09-347-064-10

Sequence 10, Application US/09347064A  
 Patent No. US20020045208A1

GENERAL INFORMATION:  
 APPLICANT: Eck, Jürgen  
 APPLICANT: Schmidt, Arno  
 APPLICANT: Zinke, Holger  
 TITLE OF INVENTION: Recombinant Fusion Proteins Based on  
 TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum  
 FILE REFERENCE: 09282-5  
 CURRENT APPLICATION NUMBER: US/09/347,064A  
 CURRENT FILING DATE: 1999-07-02  
 EARLIER APPLICATION NUMBER: PCT/EP98/00009  
 EARLIER FILING DATE: 1998-01-02  
 EARLIER APPLICATION NUMBER: EP 97 10 0012.0  
 EARLIER FILING DATE: 1997-01-02  
 NUMBER OF SEQ ID NOS: 38  
 SOFTWARE: Patentln Ver. 2.1  
 SEQ ID NO: 10  
 LENGTH: 263  
 TYPE: PRT  
 ORGANISM: Viscum album  
 US-09-347-064-10

Query Match 29.7%; Score 906; DB 9; Length 263;  
 Best Local Similarity 63.2%; Pred. No. 1.9e-80;  
 Matches 163; Conservative 32; Mismatches 63; Indels 0; Gaps 0;

QY 318 CNDPEYRIYVGRNGLCVDYRGRHNGNAIOLAMCKSNTPDANOLMTLKRONTIRSNKC 377  
 DB 5 CSASEYRIYVGRNGLCVDYRGRHNGNAIOLAMCKSNTPDANOLMTLKRONTIRSNKC 64  
 QY 378 LTTGYSPGVYMIYDCNATATATRMQWMDGTLINRSSLVLAATSGSGTTLTYOTN 437  
 DB 65 LTTGYTAGYVIMIDCNATATATRMQWMDGTLINRSSLVLAATSGSGTTLTYOTN 124  
 QY 438 IYAVSQWLPNTNTPFTVIGLYGLQIANSQGVWIEDCSSEKAEQOMALYADGSTRP 497  
 DB 125 DYTIGQWLAGNDTAPREVITIGFRDLCHESNGSGVWETCVSSQKNRMALYADGSTRP 184  
 QY 498 QONRNCGLTSDSNIRKTYVKKILSCGPASSGQWMMFKNDGTLINLYSGLVYDVRASDPGLK 557  
 DB 185 KQNDQCLTCGRDVSSTVINIVSCASGSGQWMMFKNDGTLINLYSGLVYDVRASDPGLK 244  
 QY 558 QIILYPLHGDPNQIMPL 575





Db 229 DAGVTSNIALLNLR 243

## RESULT 7

US-10-127-890-6

Sequence 6, Application US/10127890

Publication No. US20030166196A1

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

Carroll, Stephen F.

Studnika, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

City: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-May-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-May-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-May-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-Dec-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-Jun-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-Nov-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 247 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-127-890-6

Query Match 13.6%; Score 416; DB 12; Length 247;

Best Local Similarity 37.3%; Pred. No. 2.3e-32;

Matches 95; Conservative 59; Mismatches 83; Indels 18; Gaps 8;

44 INFTAGATVOSTYNTIRAVRGRLTGADVHREIPVLPNRVGLPINORFIVELSNHML 103

2 VSFRLSGATSSSYGVFISNLRKALPNERKL-YDIPLL-RSSLPSSQKRALHLINLYADE 58

104 SVTLADVTNNAVYVGRAGNSAVFFHPDNOEDA-EAITHLFTDVONRYTFAFGNVDRL 162

59 TISVAIDVTNVIYIMGRAGDTSYFF--NEASATFEAKYVFKDMARKVTLPLYSNGYERLQ 115

163 QLAGNLRNENIELGNGPLEEASIALYVYSGTGOLPTLARSFICIONISEARFQYIEGE 222

116 TRACKIRNRNIPGLPALDSATITFLFYNNAN-----SAASALMYLIQSTSEARVKTIEQ 170

223 MKTRIRYNRKSRAPDPSVITLLENSWGRSTAIO--ESNOGAFASPIQORNGSKFSY--Y 278

171 IGRKV--DKTFPLPSAIITLSENSWSALSKOIQIASTNNGQFESPVLLINNOVATITNV 228

279 DVSILIPITIALMYR 293

229 DAGVTSNIALLNLR 243

## RESULT 8

US-10-280-679B-4

Sequence 4, Application US/10280679B

Publication No. US20030150019A1

GENERAL INFORMATION:

APPLICANT: Large Scale Biology Corporation

TITLE OF INVENTION: Monopartite RNA Virus Transformation Vectors

FILE REFERENCE: LSBC-0109-US03

CURRENT APPLICATION NUMBER: US/10/280,679B

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: 09/557,941

PRIOR FILING DATE: 2000-04-24

PRIOR APPLICATION NUMBER: 08/484,341

PRIOR FILING DATE: 1995-06-07

PRIOR APPLICATION NUMBER: 07/923,692

PRIOR FILING DATE: 1992-07-31

PRIOR APPLICATION NUMBER: 07/600,244

PRIOR FILING DATE: 1990-10-22

PRIOR APPLICATION NUMBER: 07/641,617

PRIOR FILING DATE: 1991-01-16

PRIOR APPLICATION NUMBER: 07/737,899

PRIOR FILING DATE: 1991-07-26

PRIOR APPLICATION NUMBER: 07/739,143

PRIOR FILING DATE: 1991-08-01

PRIOR APPLICATION NUMBER: 07/310,881

PRIOR FILING DATE: 1989-02-17

PRIOR APPLICATION NUMBER: 07/160,766

PRIOR FILING DATE: 1988-02-26

PRIOR APPLICATION NUMBER: 07/160,771

PRIOR FILING DATE: 1988-02-26

Remaining Prior Application data removed - See file Wrapper or PALM.

NUMBER OF SEQ ID NOS: 11

SOFTWARE: Patent in version 3.1

SEQ ID NO: 4

LENGTH: 289

TYPE: PRT

ORGANISM: Chinese cucumber protein alpha-trichosanthin

US-10-280-679B-4

Query Match 13.6%; Score 416; DB 12; Length 289;

Best Local Similarity 36.9%; Pred. No. 3e-32;

Matches 94; Conservative 60; Mismatches 83; Indels 18; Gaps 8;

44 INFTAGATVOSTYNTIRAVRGRLTGADVHREIPVLPNRVGLPINORFIVELSNHML 103

25 VSFRLSGATSSSYGVFISNLRKALPNERKL-YDIPLL-RSSLPSSQKRALHLINLYADE 81

104 SVTLADVTNNAVYVGRAGNSAVFFHPDNOEDA-EAITHLFTDVONRYTFAFGNVDRL 162

82 TISVAIDVTNVIYIMGRAGDTSYFF--NEASATFEAKYVFKDMARKVTLPLYSNGYERLQ 138

163 QLAGNLRNENIELGNGPLEEASIALYVYSGTGOLPTLARSFICIONISEARFQYIEGE 222

139 TANGKIRNRNIPGLPALDSATITFLFYNNAN-----SAASALMYLIQSTSEARVKTIEQ 193

223 MKTRIRYNRKSRAPDPSVITLLENSWGRSTAIO--ESNOGAFASPIQORNGSKFSY--Y 278

194 IGRKV--DKTFPLPSAIITLSENSWSALSKOIQIASTNNGQFESPVLLINNOVATITNV 251

279 DVSILIPITIALMYR 293

RESULT 10  
 US-09-347-064-8  
 : Sequence 8, Application US/09347064A  
 : Patent No. US20020045208A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Eck, Jürgen  
 : APPLICANT: Schmidt, Arno  
 : APPLICANT: Zinke, Holger  
 : TITLE OF INVENTION: Recombinant Fusion Proteins Based on  
 : TITLE OF INVENTION: Rldosome-Inactivating Proteins of the Mistletoe Viscum  
 : TITLE OF INVENTION: album  
 : FILE REFERENCE: 09282-5  
 : CURRENT APPLICATION NUMBER: US/09/347, 064A  
 : CURRENT FILING DATE: 1999-07-02  
 : EARLIER APPLICATION NUMBER: PCT/EP98/00009  
 : EARLIER FILING DATE: 1998-01-02  
 : EARLIER APPLICATION NUMBER: EP 97 10 0012.0  
 : EARLIER FILING DATE: 1997-01-02  
 : NUMBER OF SEQ ID NOS: 38  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 8  
 : LENGTH: 252  
 : TYPE: PRT  
 : ORGANISM: Viscum album

OY 46 TGAATVQSTNTIAVRGRLLTTGADVRRIETVLPENRGPLN--QRITVELSNHAEVSY 105  
 10 TQQTGGEEEFREITLLRDYVSSGS-FNSEIPL-L-RQSTIPVSDAQRFVLELTNCGDSI 67  
 Db 106 TTAIDVTNAAVYVYGRAGNSAFAFFHDNDQDEAATLHETFTVOVRNRYPAFCGNTDRLLEOLA 165  
 OY 68 TTAIDVTNLYVAVAYAGGQOSTFLR-DAPRGE--TLHFTGT-TRSSLPFGASTPDLERTA 123  
 Db 166 GNLRENIELGNPLEEALISALYYYSTGGTOLPTLARSFIICIGIMISEANFOYIEGEMRT 225  
 OY 124 GH-RDQIPGLDILQISVTALRF---PGGSTRGTQARSILLIQIMISPAANFNILMRARQ 179  
 Db 226 RRTYRRRSRAPPSPVTLLENMGRLSTAIQESNOCAGASPTQIQRNMGKFSYVDASILP 285  
 OY 180 YNNSGASFLPDVMYMELETSWGGOSTVOHSTDCVFNPNPTRLALIPGNEVTLTNVADVYA 239

OY 286 IIALMYRCAPP 298  
DB 240 SLAIMLFCGERP 252

## RESULT 12

US-10-127-890-111  
Sequence 111, Application US/10127890  
Publication No. US2003016196A1

## GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnicka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-May-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-May-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-May-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-Dec-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-Jun-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-Nov-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 111:

SEQUENCE CHARACTERISTICS:

LENGTH: 251 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 111:

US-10-127-890-111

Query Match 11.7%; Score 358; DB 12; Length 251;

Best Local Similarity 36.9%; Pred. No. 1.2e-26;

Matches 96; Conservative 45; Mismatches 99; Indels 20; Gaps 10;

OY 44 INETAGATVOSTYNTFTAVRGRLTTGADVRAHEIPVLPNRVGLPLNQRFITVELSNRAEL 103

DB 5 VSFSTGCAFTYTYVVFLELRKIKPEGN-SHGIPPLRKADDP-GKAFVIALSNDNGQ 62

OY 104 SVTLADYTNVAVVGVRAGNSAFYFH--PDNQDEADATHTLFTD-VQNRFFAFGANYDR 160

DB 63 LAELADVTYVVGIVGRNRSYFPKADP-----AAEGLFKNTIKR--LHFGSTPS 115

OY 161 LEQLAGN--LRENIELGNGPLEFAISALYXXSTGTOLPTLARSFIICMISEARFQY 218

DB 116 LE--GEKAYRETTDLGTEPLRIGIKLDEADAINYPTLASSLLVVIQVSEARFTF 172

OY 219 IEGEMRTIRIRYRRSAPPSPVITLNSWGRSTAIQESN-QGAFASPIQORRGSFVS 277

DB 173 IENQIRN--NFQRIRPANNTISLENKWKILSFQIRSGANGMFSFAVEIERANGKRYV 230

OY 278 YDVSILITPIALMYRCAPP 297

DB 231 TAVDQVKRKIALLRKFKCDP 250

## RESULT 13

US-10-127-890-102

Sequence 102, Application US/10127890

Publication No. US2003016196A1

GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnicka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-May-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-May-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-May-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-Dec-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-Jun-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-Nov-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:

LENGTH: 251 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 102:

US-10-127-890-102

Query Match 11.6%; Score 354; DB 12; Length 251;

Best Local Similarity 36.9%; Pred. No. 2.9e-26;



SEQUENCE DESCRIPTION: SEQ ID NO: 251;  
US-09-765-527-251

Query Match 11.4%; Score 348.5; DB 9; Length 332;

Best Local Similarity 35.0%; Pred. No. 1.6e-25;

Matches 104; Conservative 50; Mismatches 118; Indels 25; Gaps 12;

```
QY 44 INFETAGATVOSTYNTFRAYRGRLTTGADYRHEIPVLPNRVGLPINORFILVELSNHAEI 103
Db 27 VSFSTKGATYITVYNFLNELRVKLRKPEGN-SHGIPILRKCDDE-GKCFVLVALSNDNGQ 84
QY 104 SVTLALDVTNAYVYVGYRAGNSAYFEH--PDNQEDAEALTHLFTD-VONRYTFAGGNYDR 160
Db 85 LAETAIIDVTSYVYVGYRNRSTYFFKDAPD-----AAVEGLFKNTIKTR--LHFGGYPS 137
QY 161 LEOLAGN--LRENIELGNPLEEAISALYYSTGTQLPTLARSFIICIMISEARFQY 218
Db 138 LE--GEKAYRETDGLIEPLRIGIKKIDENALIDNYKPTETIASLVLVQMVSEAAAFTE 194
QY 219 IEGEMRTIRIKNRSAEDPSYITLENMGRISTAIQESN-QGAFASPIQLQRNGSKFSY 277
Db 195 IENQIRN--NFQRIIPANNNTISLENKWKLSFQIRTSGANMFSSEAVELERANGKKYYV 252
QY 278 YDVSILIPILALMYRCAPPPS---QFSLLIRPVVPPNFNADVCMDPEPIVRIYGR 330
Db 253 TAVDQVVKPKIALLKFFVDKDPKSAACHHNASRVARMASDEPPSMCAMALDPI-KISGK 308
```

Search completed: September 16, 2003, 12:09:54  
Job time : 60.5688 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:40:32; Search time 37.3567 Seconds  
(without alignments)  
1482.817 Million cell updates/sec

Title: US-10-083-336a-1

Perfect score: 3051  
Sequence: 1 MKPGNTIVIMVAATWLC.....KQIIYPLHGDPMQIMPLF 576

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3051	100.0	576	1	RLCSD
2	2668.5	87.5	564	1	ricin D precursor
3	1341.5	44.0	263	2	agglutinin precurs
4	1334.5	43.7	562	2	ricin E - castor b
5	1325.5	43.4	528	2	abrin-c precursor
6	1314.5	43.1	528	1	abrin-d precursor
7	1314	43.1	527	2	abrin-a precursor
8	947.5	31.1	570	2	agglutinin I precu
9	907.5	29.7	264	2	agglutinin I
10	897	29.4	251	2	mistletoe lectin I
11	451	14.8	289	1	lectin-I B chain -
12	421	13.6	289	1	abrin (clone 7.2)
13	414	13.6	289	2	RNA N-glycosidase
14	414	13.6	289	2	karasurin - Mongol
15	413	13.5	247	2	karasurin C - Tric
16	398.5	13.1	278	2	karasurin-B - Tric
17	397	13.0	254	2	mistletoe lectin I
18	386	12.7	316	2	RNA N-glycosidase
19	370.5	12.1	245	2	RNA N-glycosidase
20	369	12.1	250	2	RNA N-glycosidase
21	354.5	11.6	286	1	RNA N-glycosidase
22	352.5	11.6	277	2	RNA N-glycosidase
23	348	11.4	286	2	RNA N-glycosidase
24	341	11.2	294	2	RNA N-glycosidase
25	337.5	11.1	294	2	RNA N-glycosidase
26	312	10.2	313	2	RNA N-glycosidase
27	295.5	9.7	261	2	antiviral protein
28	233	7.6	272	2	betavulgin - beet
29	221.5	7.3	278	2	RNA N-glycosidase

30	196	6.4	289	2	T12573
31	195.5	6.4	106	2	B39761
32	175.5	5.8	310	2	S46239
33	166	5.4	293	2	S17519
34	164	5.4	292	1	R10H62
35	160.5	5.3	253	2	S28539
36	160	5.2	377	2	JC7535
37	159	5.2	283	2	S05205
38	157	5.1	477	1	J50589
39	156.5	5.1	253	2	S28541
40	156.5	5.1	253	2	S28542
41	153	5.0	383	2	T34603
42	152.5	5.0	253	2	S29931
43	146	4.8	275	2	S33631
44	146	4.8	280	1	RLBH
45	144.5	4.7	475	2	T35697

## ALIGNMENTS

## RESULT 1

ricin D precursor - castor bean

N:Contains: RNA N-glycosidase (EC 3.2.2.22)

C:Species: Ricinus communis (castor bean)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999

C:Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903

R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.F

Nucleic Acids Res. 13, 8019-8033, 1985

A:Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.

A:Reference number: A24041; MUID:86067214; PMID:2999712

A:Accession: A24041

A:Molecule type: DNA

A:Residues: 1-576 <HAL>

A:Cross-references: GB:X03179; NID:g21082; PIDN:CAA26939.1; PID:g21083

R:Tregeat, J.W.; Roberts, L.M.

Plant Mol. Biol. 18, 515-525, 1992

A:Title: The lectin gene family of Ricinus communis: cloning of a functional ricin ge

A:Reference number: S20513; MUID:92163016; PMID:1371405

A:Accession: S20513

A:Molecule type: DNA

A:Residues: 1-576 <TRF>

A:Cross-references: EMBL:X52908; NID:g21084; PIDN:CAA37095.1; PID:g21085

R:Lamb, F.T.; Roberts, L.M.; Lord, J.M.

Eur. J. Biochem. 148, 265-270, 1985

A:Title: Nucleotide sequence of cloned cDNA coding for preproricin.

A:Reference number: A24614; MUID:85179479; PMID:3838723

A:Accession: A24614

A:Molecule type: mRNA

A:Residues: 12-75, 'D', '77-550, 'R', '552-576 <LAM>

A:Cross-references: GB:X03388; NID:g21077; PIDN:CAA26230.1; PID:g21078

R:Yoshitake, S.; Funatsu, G.; Funatsu, M.

Agric. Biol. Chem. 42, 1267-1274, 1978

A:Title: Isolation and sequences of peptic peptides, and the complete sequence of the

A:Reference number: A03372

A:Accession: A03372

A:Molecule type: protein

A:Residues: 36-97, 'Q', '99-109, 'S', '111-269, 'D', '272-283, 'L', '285-288, '290-302 <YOS>

A>Note: this paper cites the others in the series providing experimental details for

R:Ataki, T.; Funatsu, G.

FEBS Lett. 191, 121-124, 1985

A:Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptop

A:Reference number: A24010

A:Accession: A24010

A:Molecule type: protein

A:Residues: 315-383, 'PS', '386-576 <ARA>

R:Funatsu, G.; Kimura, M.; Funatsu, M.

Agric. Biol. Chem. 43, 2221-2224, 1979

A:Title: Primary structure of Ala chain of ricin D.

A:Reference number: A03374

A:Accession: A03374

A:Molecule type: protein

A:Residues: 315-335, 'N', 337-342, 'NH', 345-362, 364-383, 'PS', 386-399, 'T', 401, 'D', 403, 'E', 405, 521, 'E', 529-564, 'W', 566, 'H', 567-570, 'L', 573-574, 'F' <FUN>  
A:Note: this paper, one of a series, summarizes the experimental details for the determination of the structure of ricin A-chain and implications for the mechanism of action of ricin A-chain and B-chain.  
R:Ready, M.P., Kim, X.; Robertus, J.D.  
Proteins 10, 270-278, 1991  
A:Title: Site-directed mutagenesis of ricin A-chain and implications for the mechanism of action of ricin A-chain and B-chain.  
A:Reference number: A48237; PMID:91352006; PMID:1881883  
A:Contents: annotation; active site  
R:Robertus, J.D.  
Proteins 10, 260-269, 1991  
A:Title: Structure of ricin B-chain at 2.5 angstrom resolution.  
A:Reference number: A48238; PMID:91352005; PMID:1881882  
A:Contents: annotation; X-ray crystallography, 2.5 angstroms  
R:Katzin, B.J.; Collins, E.J.; Robertus, J.D.  
Proteins 10, 251-259, 1991  
A:Title: Structure of ricin A-chain at 2.5 angstroms.  
A:Reference number: A48239; PMID:91352004; PMID:1881881  
A:Contents: annotation; X-ray crystallography, 2.5 angstroms  
A:Comment: The functional molecule is a disulfide-linked dimer of A and B chains, which act together to inhibit protein synthesis. It inactivates the 60S ribosomal subunit of the cell of the A chain; B chains are also responsible for cell agglutination (lectin activity).  
C:Comment: This protein is cytotoxic and very poisonous to animals.  
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed  
F:1-35/Domain: signal sequence #status predicted <SIG>  
F:35-302/Product: ricin D chain A #status experimental <ACH>  
F:35-302/Product: ricin D chain B #status experimental <ACH>  
F:46-293/Domain: RNA N-glycosidase homology <RNG>  
F:315-576/Product: ricin D chain B #status experimental <BCH>  
F:331-373, 374-414, 417-455, 462-497, 501-540, 543-576/Region: 40-residue repeats  
F:409, 449/Binding site: carboxylate (Asn) (covalent) #status experimental  
F:115, 158, 243, 244/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
F:212/Active site: Glu #status experimental  
F:215/Active site: Arg #status predicted  
F:294-318, 334-353, 377-394, 465-478, 504-521/Disulfide bonds: #status experimental  
F:336, 349, 360/Binding site: N-acetylgalactosamine (Asp, Glu, Asn) #status experimental  
F:548, 569/Binding site: N-acetylgalactosamine (Asp, Asn) #status experimental

Query Match 100.0%; Score 3051; DB 1; Length 576;  
Best Local Similarity 100.0%; Pred. No. 1.7e-224;  
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPGNTIVIMVAWVATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTAGATVOSTYNTF 60  
DB 1 MKPGNTIVIMVAWVATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTAGATVOSTYNTF 60  
QY 61 RAVGRITGADVHETIPVLPNRYGLPINORFIIIVELSNHAELSVTLADVTNAYVGYR 120  
DB 61 RAVGRITGADVHETIPVLPNRYGLPINORFIIIVELSNHAELSVTLADVTNAYVGYR 120  
QY 121 AGNSAFEPHPDNOEDAEITLFTDVONRYTFAFGNYDRLEQLAGMLRENIELGNGPLE 180  
DB 121 AGNSAFEPHPDNOEDAEITLFTDVONRYTFAFGNYDRLEQLAGMLRENIELGNGPLE 180  
QY 181 EAIISALYYSTGGQOLPLTARSFTICOMISEARFOYIEGEMTRIRYRRSAPDSVI 240  
DB 181 EAIISALYYSTGGQOLPLTARSFTICOMISEARFOYIEGEMTRIRYRRSAPDSVI 240  
QY 241 TLENSWGRSTAIQESNOCASPIQOLRRNGSKFSYDVSIILPIIALMYRCAPPSS 300  
DB 241 TLENSWGRSTAIQESNOCASPIQOLRRNGSKFSYDVSIILPIIALMYRCAPPSS 300  
QY 301 QFSLLIRVPVFNADVCMDEPIYRIYGRNGLCYDADGFRHNGNAIOLMPCKSNIDAN 360  
DB 301 QFSLLIRVPVFNADVCMDEPIYRIYGRNGLCYDADGFRHNGNAIOLMPCKSNIDAN 360  
QY 361 QLTWTKRNTTIRSNKCLTITGYSPGVVMTYDCATATDARMOIMNGTILIRPSLY 420  
DB 361 QLTWTKRNTTIRSNKCLTITGYSPGVVMTYDCATATDARMOIMNGTILIRPSLY 420  
QY 421 LAATSGNSGTLTVOJNIIYAVSQGLPTNNTOPEVTITVGLXGLCLQANSQVWIEDCS 480  
DB 421 LAATSGNSGTLTVOJNIIYAVSQGLPTNNTOPEVTITVGLXGLCLQANSQVWIEDCS 480

QY 481 EKAQOQWALYADSGIRPOQRNDCLTSDNSNIRETVKILISCGPASSGQRMFKNDGTILN 540  
DB 481 EKAQOQWALYADSGIRPOQRNDCLTSDNSNIRETVKILISCGPASSGQRMFKNDGTILN 540  
QY 541 LYSGLVLDVRAVSPSLKQIITLPLHGDPPNOIWLPLF 576  
DB 541 LYSGLVLDVRAVSPSLKQIITLPLHGDPPNOIWLPLF 576

RESULT 2  
R:CSAG  
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
C:Species: Ricinus communis (castor bean)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
C:Accession: A24261; A24210  
R:Robertus, J.D.; Lamb, J.F.; Pappin, D.J.C.; Lord, J.M.  
J. Biol. Chem. 260, 15682-15686, 1985  
A:Title: The primary sequence of Ricinus communis agglutinin. Comparison with ricin.  
A:Reference number: A24261; PMID:86059449; PMID:2999130  
A:Accession: A24261  
A:Molecule type: mRNA  
A:Residues: 1-564 <ROB>  
A:Cross-references: GB:M12089; NID:g169700; PIDN:AAA3869.1; PID:g169701  
R:Araki, T.; Yoshioke, Y.; Funatsu, G.  
Biochim. Biophys. Acta 872, 277-285, 1986  
A:Title: The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin.  
A:Reference number: A24210  
A:Accession: A24210  
A:Molecule type: protein  
A:Residues: 303-325, 'F', 327-330, 'T', 332-361, 'D', 363-373, 'G', 375-403, 'T', 405-551, 'V', 553-564, 566-582, 584-585, 587-590, 592-599, 601-602, 604-605, 607-608, 610-611, 613-614, 616-617, 619-620, 622-623, 625-626, 628-629, 631-632, 634-635, 637-638, 640-641, 643-644, 646-647, 649-650, 652-653, 655-656, 658-659, 661-662, 664-665, 667-668, 670-671, 673-674, 676-677, 679-680, 682-683, 685-686, 688-689, 691-692, 694-695, 697-698, 700-701, 703-704, 706-707, 709-710, 712-713, 715-716, 718-719, 721-722, 724-725, 727-728, 730-731, 733-734, 736-737, 739-740, 742-743, 745-746, 748-749, 751-752, 754-755, 757-758, 760-761, 763-764, 766-767, 769-770, 772-773, 775-776, 778-779, 781-782, 784-785, 787-788, 790-791, 793-794, 796-797, 799-800, 802-803, 805-806, 808-809, 811-812, 814-815, 817-818, 820-821, 823-824, 826-827, 829-830, 832-833, 835-836, 838-839, 841-842, 844-845, 847-848, 850-851, 853-854, 856-857, 859-860, 862-863, 865-866, 868-869, 871-872, 874-875, 877-878, 880-881, 883-884, 886-887, 889-890, 892-893, 895-896, 898-899, 901-902, 904-905, 907-908, 910-911, 913-914, 916-917, 919-920, 922-923, 925-926, 928-929, 931-932, 934-935, 937-938, 940-941, 943-944, 946-947, 949-950, 952-953, 955-956, 958-959, 961-962, 964-965, 967-968, 970-971, 973-974, 976-977, 979-980, 982-983, 985-986, 988-989, 991-992, 994-995, 997-998, 1000-1001, 1003-1004, 1006-1007, 1009-1010, 1012-1013, 1015-1016, 1018-1019, 1021-1022, 1024-1025, 1027-1028, 1030-1031, 1033-1034, 1036-1037, 1039-1040, 1042-1043, 1045-1046, 1048-1049, 1051-1052, 1054-1055, 1057-1058, 1060-1061, 1063-1064, 1066-1067, 1069-1070, 1072-1073, 1075-1076, 1078-1079, 1081-1082, 1084-1085, 1087-1088, 1090-1091, 1093-1094, 1096-1097, 1099-1100, 1102-1103, 1105-1106, 1108-1109, 1111-1112, 1114-1115, 1117-1118, 1120-1121, 1123-1124, 1126-1127, 1129-1130, 1132-1133, 1135-1136, 1138-1139, 1141-1142, 1144-1145, 1147-1148, 1150-1151, 1153-1154, 1156-1157, 1159-1160, 1162-1163, 1165-1166, 1168-1169, 1171-1172, 1174-1175, 1177-1178, 1180-1181, 1183-1184, 1186-1187, 1189-1190, 1192-1193, 1195-1196, 1198-1199, 1200-1201, 1203-1204, 1206-1207, 1209-1210, 1212-1213, 1215-1216, 1218-1219, 1221-1222, 1224-1225, 1227-1228, 1230-1231, 1233-1234, 1236-1237, 1239-1240, 1242-1243, 1245-1246, 1248-1249, 1251-1252, 1254-1255, 1257-1258, 1260-1261, 1263-1264, 1266-1267, 1269-1270, 1272-1273, 1275-1276, 1278-1279, 1281-1282, 1284-1285, 1287-1288, 1290-1291, 1293-1294, 1296-1297, 1299-1300, 1302-1303, 1305-1306, 1308-1309, 1311-1312, 1314-1315, 1317-1318, 1320-1321, 1323-1324, 1326-1327, 1329-1330, 1332-1333, 1335-1336, 1338-1339, 1341-1342, 1344-1345, 1347-1348, 1350-1351, 1353-1354, 1356-1357, 1359-1360, 1362-1363, 1365-1366, 1368-1369, 1371-1372, 1374-1375, 1377-1378, 1380-1381, 1383-1384, 1386-1387, 1389-1390, 1392-1393, 1395-1396, 1398-1399, 1401-1402, 1404-1405, 1407-1408, 1410-1411, 1413-1414, 1416-1417, 1419-1420, 1422-1423, 1425-1426, 1428-1429, 1431-1432, 1434-1435, 1437-1438, 1440-1441, 1443-1444, 1446-1447, 1449-1450, 1452-1453, 1455-1456, 1458-1459, 1461-1462, 1464-1465, 1467-1468, 1470-1471, 1473-1474, 1476-1477, 1479-1480, 1482-1483, 1485-1486, 1488-1489, 1491-1492, 1494-1495, 1497-1498, 1499-1500, 1502-1503, 1505-1506, 1508-1509, 1511-1512, 1514-1515, 1517-1518, 1520-1521, 1523-1524, 1526-1527, 1529-1530, 1532-1533, 1535-1536, 1538-1539, 1541-1542, 1544-1545, 1547-1548, 1550-1551, 1553-1554, 1556-1557, 1559-1560, 1562-1563, 1565-1566, 1568-1569, 1571-1572, 1574-1575, 1577-1578, 1580-1581, 1583-1584, 1586-1587, 1589-1590, 1592-1593, 1595-1596, 1598-1599, 1601-1602, 1604-1605, 1607-1608, 1610-1611, 1613-1614, 1616-1617, 1619-1620, 1622-1623, 1625-1626, 1628-1629, 1631-1632, 1634-1635, 1637-1638, 1640-1641, 1643-1644, 1646-1647, 1649-1650, 1652-1653, 1655-1656, 1658-1659, 1661-1662, 1664-1665, 1667-1668, 1670-1671, 1673-1674, 1676-1677, 1679-1680, 1682-1683, 1685-1686, 1688-1689, 1691-1692, 1694-1695, 1697-1698, 1699-1700, 1702-1703, 1705-1706, 1708-1709, 1711-1712, 1714-1715, 1717-1718, 1720-1721, 1723-1724, 1726-1727, 1729-1730, 1732-1733, 1735-1736, 1738-1739, 1741-1742, 1744-1745, 1747-1748, 1750-1751, 1753-1754, 1756-1757, 1759-1760, 1762-1763, 1765-1766, 1768-1769, 1771-1772, 1774-1775, 1777-1778, 1780-1781, 1783-1784, 1786-1787, 1789-1790, 1792-1793, 1795-1796, 1798-1799, 1801-1802, 1804-1805, 1807-1808, 1810-1811, 1812-1813, 1815-1816, 1818-1819, 1821-1822, 1824-1825, 1827-1828, 1830-1831, 1833-1834, 1836-1837, 1839-1840, 1842-1843, 1845-1846, 1848-1849, 1851-1852, 1854-1855, 1857-1858, 1860-1861, 1863-1864, 1866-1867, 1869-1870, 1872-1873, 1875-1876, 1878-1879, 1881-1882, 1884-1885, 1887-1888, 1890-1891, 1893-1894, 1896-1897, 1899-1900, 1902-1903, 1905-1906, 1908-1909, 1911-1912, 1914-1915, 1917-1918, 1920-1921, 1923-1924, 1926-1927, 1929-1930, 1932-1933, 1935-1936, 1938-1939, 1941-1942, 1944-1945, 1947-1948, 1950-1951, 1953-1954, 1956-1957, 1959-1960, 1962-1963, 1965-1966, 1968-1969, 1971-1972, 1974-1975, 1977-1978, 1980-1981, 1983-1984, 1986-1987, 1989-1990, 1992-1993, 1995-1996, 1998-1999, 2001-2002, 2004-2005, 2007-2008, 2010-2011, 2013-2014, 2016-2017, 2019-2020, 2022-2023, 2025-2026, 2028-2029, 2031-2032, 2034-2035, 2037-2038, 2040-2041, 2043-2044, 2046-2047, 2049-2050, 2052-2053, 2055-2056, 2058-2059, 2061-2062, 2064-2065, 2067-2068, 2070-2071, 2073-2074, 2076-2077, 2079-2080, 2082-2083, 2085-2086, 2088-2089, 2091-2092, 2094-2095, 2097-2098, 2099-2100, 2102-2103, 2105-2106, 2108-2109, 2111-2112, 2114-2115, 2117-2118, 2120-2121, 2123-2124, 2126-2127, 2129-2130, 2132-2133, 2135-2136, 2138-2139, 2141-2142, 2144-2145, 2147-2148, 2150-2151, 2153-2154, 2156-2157, 2159-2160, 2162-2163, 2165-2166, 2168-2169, 2171-2172, 2174-2175, 2177-2178, 2180-2181, 2183-2184, 2186-2187, 2189-2190, 2192-2193, 2195-2196, 2198-2199, 2201-2202, 2204-2205, 2207-2208, 2210-2211, 2213-2214, 2216-2217, 2219-2220, 2222-2223, 2225-2226, 2228-2229, 2231-2232, 2234-2235, 2237-2238, 2240-2241, 2243-2244, 2246-2247, 2249-2250, 2252-2253, 2255-2256, 2258-2259, 2261-2262, 2264-2265, 2267-2268, 2270-2271, 2273-2274, 2276-2277, 2279-2280, 2282-2283, 2285-2286, 2287-2288, 2290-2291, 2293-2294, 2296-2297, 2299-2300, 2302-2303, 2305-2306, 2308-2309, 2311-2312, 2314-2315, 2317-2318, 2320-2321, 2323-2324, 2326-2327, 2329-2330, 2332-2333, 2335-2336, 2338-2339, 2341-2342, 2344-2345, 2347-2348, 2350-2351, 2353-2354, 2356-2357, 2359-2360, 2362-2363, 2365-2366, 2368-2369, 2371-2372, 2374-2375, 2377-2378, 2380-2381, 2383-2384, 2386-2387, 2389-2390, 2392-2393, 2395-2396, 2398-2399, 2401-2402, 2404-2405, 2407-2408, 2410-2411, 2413-2414, 2416-2417, 2419-2420, 2422-2423, 2425-2426, 2428-2429, 2431-2432, 2434-2435, 2437-2438, 2440-2441, 2443-2444, 2446-2447, 2449-2450, 2452-2453, 2455-2456, 2458-2459, 2461-2462, 2464-2465, 2467-2468, 2470-2471, 2473-2474, 2476-2477, 2479-2480, 2482-2483, 2485-2486, 2488-2489, 2491-2492, 2494-2495, 2497-2498, 2499-2500, 2502-2503, 2505-2506, 2508-2509, 2511-2512, 2514-2515, 2517-2518, 2520-2521, 2523-2524, 2526-2527, 2529-2530, 2532-2533, 2535-2536, 2538-2539, 2541-2542, 2544-2545, 2547-2548, 2550-2551, 2553-2554, 2556-2557, 2559-2560, 2562-2563, 2565-2566, 2568-2569, 2571-2572, 2574-2575, 2577-2578, 2580-2581, 2583-2584, 2586-2587, 2589-2590, 2592-2593, 2595-2596, 2598-2599, 2601-2602, 2604-2605, 2607-2608, 2610-2611, 2613-2614, 2616-2617, 2619-2620, 2622-2623, 2625-2626, 2628-2629, 2631-2632, 2634-2635, 2637-2638, 2640-2641, 2643-2644, 2646-2647, 2649-2650, 2652-2653, 2655-2656, 2658-2659, 2661-2662, 2664-2665, 2667-2668, 2670-2671, 2673-2674, 2676-2677, 2679-2680, 2682-2683, 2685-2686, 2688-2689, 2691-2692, 2694-2695, 2697-2698, 2699-2700, 2702-2703, 2705-2706, 2708-2709, 2711-2712, 2714-2715, 2717-2718, 2720-2721, 2723-2724, 2726-2727, 2729-2730, 2732-2733, 2735-2736, 2738-2739, 2741-2742, 2744-2745, 2747-2748, 2750-2751, 2753-2754, 2756-2757, 2759-2760, 2762-2763, 2765-2766, 2768-2769, 2771-2772, 2774-2775, 2777-2778, 2780-2781, 2783-2784, 2786-2787, 2789-2790, 2792-2793, 2795-2796, 2798-2799, 2801-2802, 2804-2805, 2807-2808, 2810-2811, 2812-2813, 2815-2816, 2818-2819, 2821-2822, 2824-2825, 2827-2828, 2830-2831, 2833-2834, 2836-2837, 2839-2840, 2842-2843, 2845-2846, 2848-2849, 2851-2852, 2854-2855, 2857-2858, 2860-2861, 2863-2864, 2866-2867, 2869-2870, 2872-2873, 2875-2876, 2878-2879, 2881-2882, 2884-2885, 2887-2888, 2890-2891, 2893-2894, 2896-2897, 2899-2900, 2902-2903, 2905-2906, 2908-2909, 2911-2912, 2914-2915, 2917-2918, 2920-2921, 2923-2924, 2926-2927, 2929-2930, 2932-2933, 2935-2936, 2938-2939, 2941-2942, 2944-2945, 2947-2948, 2950-2951, 2953-2954, 2956-2957, 2959-2960, 2962-2963, 2965-2966, 2968-2969, 2971-2972, 2974-2975, 2977-2978, 2980-2981, 2983-2984, 2986-2987, 2989-2990, 2992-2993, 2995-2996, 2998-2999, 3001-3002, 3004-3005, 3007-3008, 3010-3011, 3013-3014, 3016-3017, 3019-3020, 3022-3023, 3025-3026, 3028-3029, 3031-3032, 3034-3035, 3037-3038, 3040-3041, 3043-3044, 3046-3047, 3049-3050, 3052-3053, 3055-3056, 3058-3059, 3061-3062, 3064-3065, 3067-3068, 3070-3071, 3073-3074, 3076-3077, 3079-3080, 3082-3083, 3085-3086, 3088-3089, 3091-3092, 3094-3095, 3097-3098, 3099-3100, 3102-3103, 3105-3106, 3108-3109, 3111-3112, 3114-3115, 3117-3118, 3120-3121, 3123-3124, 3126-3127, 3129-3130, 3132-3133, 3135-3136, 3138-3139, 3141-3142, 3144-3145, 3147-3148, 3150-3151, 3153-3154, 3156-3157, 3159-3160, 3162-3163, 3165-3166, 3168-3169, 3171-3172, 3174-3175, 3177-3178, 3180-3181, 3183-3184, 3186-3187, 3189-3190, 3192-3193, 3195-3196, 3198-3199, 3201-3202, 3204-3205, 3207-3208, 3210-3211, 3213-3214, 3216-3217, 3219-3220, 3222-3223, 3225-3226, 3228-3229, 3231-3232, 3234-3235, 3237-3238, 3240-3241, 3243-3244, 3246-3247, 3249-3250, 3252-3253, 3255-3256, 3258-3259, 3261-3262, 3264-3265, 3267-3268, 3270-3271, 3273-3274, 3276-3277, 3279-3280, 3282-3283, 3285-3286, 3287-3288, 3



Db 300 NFNAVADPEPIVAVGRNGLVDYGEFFDGNPIQLMPCKSNTDMNQLMTRKDSYI 359

QY 372 RSNKGLTFTYSGPVVMIYDCNTAATDATRMOIMDNGTIINPRSLVLAATSGNSGT 431

Db 360 RSNKGLTFTYSGPVVMIYDCNTAATDATRMOIMDNGTIINPRSLVLAATSGNSGT 419

QY 432 LTVQNTIYAVSOGWLPNTNTPPVTTIYGLGLCLQANSQGVWIEDCSSEKAEQOMALYA 491

Db 420 LTVQNTIYAVSOGWLPNTNTPPVTTIYGLGLCLQANSQGVWIEDCSSEKAEQOMALYA 479

QY 492 DGSIRPOONRNDCLTSDSNIREFYVKILSCGPASSGORMFKNDGTILNLVGLVDRASD 551

Db 480 DGSIRPOONRNDCLTSDSNIREFYVKILSCGPASSGORMFKNDGTILNLVGLVDRASD 539

QY 552 SDPSLKQIILYPLHGDPMQIWLPLF 576

Db 540 SDPSLKQIILYPLHGDPMQIWLPLF 564

## RESULT 3

S06330  
ricin E - castor bean (fragment)  
N:Contains: chain B; rRNA N-glycosidase (EC 3.2.2.22) (chain A)  
C:Species: Ricinus communis (castor bean)  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 20-Aug-1999  
C:Accession: S06330; J05381  
R:Ladin, B.F.; Murray, E.E.; Halling, A.C.; Halling, K.C.; Tilakaratne, N.; Long, G.L.;  
Plant Mol. Biol. 9, 287-295, 1987  
A:Title: Characterization of a cDNA encoding ricin E, a hybrid ricin-Richinus communis ag  
A:Reference number: S06330  
A:Accession: S06330  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-263 <LAD>  
A:Cross-references: EMBL:M17631; NID:9169714; PIDN:AAA63506.1; PID:9169715  
R:Araki, T.; Funatsu, G.  
Biochim. Biophys. Acta 911, 191-200, 1987  
A:Title: The complete amino acid sequence of the B-chain of ricin E isolated from small-  
A:Reference number: J05381; M01D:87101186; PMID:3801493  
A:Accession: J05381  
A:Molecule type: Protein  
A:Residues: 1-69; 'PS', 72-209, 'A', 211-241, 243-250, 'V', 252-263 <ARA>  
A:Experimental source: seed  
C:Superfamily: ricin: rRNA N-glycosidase homology  
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase  
F:1-263/Product: ricin E chain B #status experimental <BOH>  
F:17-59,60-100,103-141,148-183,187-226,229-263/Region: 40-residue repeats  
F:20-39,63-80,151-164,190-207/Disulfide bonds: #status predicted  
F:95,135/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 44.08; Score 1341.5; DB 2; Length 263;  
Best Local Similarity 93.58; Pred. No. 1e-94;  
Matches 246; Conservative 11; Mismatches 5; Indels 1; Gaps 1;

QY 315 ADVCMDEPIVRIYGRNGLVDYRDRGFHNGNAIOLMPCKSNTDANQMLTKRDNITRSN 374

Db 1 ADVCMDEPIVRIYGRNGLVDYRDRGFHNGNAIOLMPCKSNTDANQMLTKRDNITRSN 60

QY 375 GKCLTFTYSGPVVMIYDCNTAATDATRMOIMDNGTIINPRSLVLAATSGNSGTLLY 434

Db 61 GKCLTFTYSGPVVMIYDCNTAATDATRMOIMDNGTIINPRSLVLAATSGNSGTLLY 120

QY 435 QNTIYAVSOGWLPNTNTPPVTTIYGLGLCLQANSQGVWIEDCSSEKAEQOMALYADGS 494

Db 121 QNTIYAVSOGWLPNTNTPPVTTIYGLGLCLQANSQGVWIEDCSSEKAEQOMALYADGS 180

QY 495 IRPOONRNDCLTSDSNIREFYVKILSCGPASSGORMFKNDGTILNLVGLVDRASD 554

Db 181 IRPOONRNDCLTSDSNIREFYVKILSCGPASSGORMFKNDGTILNLVGLVDRASD 240

QY 555 -SLKQIILYPLHGDPMQIWLPLF 576

Db 241 -SLKQIILYPLHGDPMQIWLPLF 263

## RESULT 4

abrin-c precursor - Indian licorice  
S16022  
N:Contains: abrin N-glycosidase (EC 3.2.2.22)  
C:Species: Abrus precatorius (Indian licorice)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Aug-1999  
C:Accession: S16022  
R:Wood, K.A.; Lord, J.M.; Wawrzyniczak, E.J.; Platek, M.  
Eur. J. Biochem. 198, 723-732, 1991  
A:Title: Preproabrin: genomic cloning, characterisation and the expression of the A-c  
A:Reference number: S16022; M01D:91266957; PMID:2050149  
A:Accession: S16022  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-562 <MOO>  
A:Cross-references: EMBL:X55667; NID:916084; PIDN:CAA39202.1; PID:916085  
C:Comment: The A and B chains are linked by a single disulfide bond, which is essential for tox  
C:Superfamily: ricin: rRNA N-glycosidase homology  
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglyutamic a  
F:35-285/Product: abrin-c chain A #status predicted <ACH>  
F:41-280/Domain: rRNA N-glycosidase homology <RNG>  
F:295-562/Product: abrin-c chain B #status predicted <BOH>  
F:317-359,360-400,403-441,448-483,487-526,529-562/Region: 40-residue repeats  
F:35/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic  
F:108,147,229,230/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
F:198,201/Active site: Glu, Arg #status predicted  
F:224,287,395,435,436/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:281-303,320-339,363-380,451-464,490-507/Disulfide bonds: #status predicted  
F:322,346/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
F:534,555/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 43.78; Score 1334.5; DB 2; Length 562;  
Best Local Similarity 48.98; Pred. No. 1e-93;  
Matches 276; Conservative 80; Mismatches 191; Indels 17; Gaps 8;

QY 19 LCGSTSGMFTLEDNINP---KQPIINTFGATVQSTNTRAVRGTLGAYVR 74

Db 10 LCLMTGCSFALRCARITPPVATNDQVYKFTTBGASQSKOTIEAROKLNGG--LI 67

QY 75 HEIPLVLPNRVGLPIINOFILVELSNHAEISVTLADVNAVYVGRAGNSAFFHPDNOE 134

Db 68 HDIPLVLPDPTVEERRRYITVELSNERSIEVIGDYNNAVVAARASQSYFL--RDA 124

QY 135 DAEALHTLFTDQNKRTFAFGNRYDRLEQALNRENLBGLPLEAIALYVYSTGCT 194

Db 125 PASASTYLPFGTY-RYSLEFDSYGDLEHWAHQTRREISLGLALTHAIS--FLRS GAS 180

QY 195 QLPTLARSFIICTOMISEARPOYTEGEMRTIRYRNRSAPDPSPVITLNSMGRSLTAIQ 254

Db 181 NDEEARKTLIVITQMASEAKRYTISNRKGVSIKGTAFQPPAPALSLNNMNDLSGGVQ 240

QY 255 ESNQAFASPILOLORRNSKESYDVS--ILPIIALMYRYCARPPSSQSFSLIRPVVNF 313

Db 241 QSVQDTPFPNNVILSSINRQPVVDSLSPHTYAVLALMLFVCPNPANQSPPLINSIVE-- 298

QY 314 NADVCMDE-EPYVRIYGRNGLVDYRDRGFHNGNAIOLMPCKSNTDANQMLTKRDNITR 372

Db 299 ESKTSSREYEPVRIYGRNGLVDYRDRGFHNGNAIOLMPCKSNTDANQMLTKRDNITR 358

QY 373 SNGKCLTFTYSGPVVMIYDCNTAATDATRMOIMDNGTIINPRSLVLAATSGNSGTLL 432

Db 359 SNGKCLTFTYSGPVVMIYDCNTAATDATRMOIMDNGTIINPRSLVLAATSGNSGTLL 418

QY 433 TVQNTIYAVSOGWLPNTNTPPVTTIYGLGLCLQANSQGVWIEDCSSEKAEQOMALYAD 492

Db 419 TVQNTIYAVSOGWLPNTNTPPVTTIYGLGLCLQANSQGVWIEDCSSEKAEQOMALYAD 478

QY 493 GSIRPOONRNDCLTSDSNIREFYVKILSCGPASSGORMFKNDGTILNLVGLVDRAS 552

Db 479 GSIRPOONRNDCLTSDSNIREFYVKILSCGPASSGORMFKNDGTILNLVGLVDRAS 538



R.Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.  
 Eur. J. Biochem. 240, 564-569, 1996  
 A:Title: Probing the domain structure of abrin-a by tryptic digestion.  
 A:Reference number: S74110; MUID:97008945; PMID:8856055  
 A:Accession: S74110  
 A:Molecule type: protein  
 A:Residues: 89-108;154-172 <LIN>  
 A:Experimental source: seed  
 A:Accession: S74111  
 A:Molecule type: protein  
 A:Residues: 262-276, 'X', 278-280;329-348;369-388;399-418 <LIN>  
 A:Experimental source: seed  
 C:Comment: Abirin-a is more toxic than ricin. The toxin consists of an A chain, which inhibits the binding of the toxin to the cell surface. The A and B chains are linked by a single disulfide bond.  
 C:Superfamily: ricin; RNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglycolytic acid  
 F:1-251/Product: abrin-a chain A #status experimental <ACH>  
 F:7-246/Domain: RNA N-glycosidase homology <RNG>  
 F:261-528/Product: abrin-a chain B #status experimental <BCH>  
 F:283-325/Product: 369-407,414-449,453-492,495-528/Region: 40-residue repeats  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:164,167/Active site: Glu, Arg #status predicted  
 F:247-269,286-305,329-346,417-430,456-473/Disulfide bonds: #status predicted  
 F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 F:361,401/Binding site: carboxylate (Asn) (covalent) #status experimental  
 F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 43.1%; Score 1314.5; DB 1; Length 528;  
 Best Local Similarity 49.7%; Pred. No. 3.2e-92;  
 Matches 267; Conservative 78; Mismatches 175; Indels 17; Gaps 8;

44 INFTAGATVQSYNTNFRAYRGRITGADVREHPIVLPNVRGPIINORFTVLSNHAEL 103  
 5 IKESTGATGTSQSYKQFTEALRELRGG--LIHIDPVLDPDTTQDENRRTVLSNDPE 62  
 104 SVTLADVTNAYVYGRAGNSAYFFH--PDNEDAEITHLFTDYNRRTFAFGNYDR 161  
 63 SIEGIDVTNAYVYVAGTOSYFLRDASSASD-----YLFPTG--DOHSLPYGYGDL 116  
 162 EQLAGNLRNIEELGNPLEEASALYYSTGQPLTLARSPFICIQMSEARROYIG 221  
 117 ERNAGHQRQIPGLQALHTGIS--FFSSGGNDNEKATLTVIOMVAEARFYISN 173  
 222 EMRTIRIRNRSAPDSVTITLNSMGRSLTAIOESNOGAFASPIQLORNGSKFSYDVS 281  
 174 RYVSIOTGTAFQPDAMISLENNMNLNGVQESYQDFPNQVTLTLNRNEVYIDSLIS 233  
 282 -LIIPIIALMYKCAPPSQSLLRPVYVFNADYCDP--EPIVRIYGRNGLCYVDRD 339  
 234 HPTVAVALALMLFVCNPNPANNOSPLIRISVE--KSKICSSREPEYRIGGRGMCYVDYD 291  
 340 GREHNGNATQIMPCCKNTDANOLMTLKRNTIRSNKCLTTYGSGVYVMYIDCQAT 399  
 292 NGHNGNRITIMKCKRLBNOLMTLKSQKTIIRSNKCLTTYGAGSTVMTYIDCQATA 351  
 400 DATRMOIWDNGTITINRSSLVLAATSGNGTTLTYOTNTYAVSOGMLPTNTPOPVY 459  
 352 EATYWEIWDNGTITINRSSLVLAATSGNGTTLTYOTNTYAVSOGMLPTNTPOPVY 411  
 460 GLYGLCLQANSGQVWIEDCSSEKAEQOMALYADGSTRPOQRNDCITSDSNIRETVYK 519  
 412 GYSGLQOMAGSVMWMAADDNKKRQOMALYTDGSTRSVOANNCTLSDHKGSGTILLM 471  
 520 SCGPASSGGRMMFKNDGTITLNLXSGVLVDYRASDPISLKOIILYPLHGDNPQIMLP 576  
 472 GCGNGMASGRWYFKNDGTITLNLXSGVLVDYRASDPISLKOIILYPLHGDNPQIMLP 528

C:Species: Abrus precatorius (Indian jicorice)  
 C:Date: 30-Sep-1993 #sequence\_revision 01-Aug-1997 #text\_change 20-Aug-1999  
 C:Accession: S32430, JCI1399  
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
 J. Mol. Biol. 229, 263-267, 1993  
 A:Title: Primary structure of three distinct isoabrin determined by cDNA sequencing.  
 A:Reference number: S32429; MUID:93132798; PMID:8421313  
 A:Accession: S32430  
 A:Molecule type: mRNA  
 A:Residues: 1-527 <HUN>  
 A:Cross-references: GB:M98345; NID:q166296; PIDN:AAA32625.1; PID:q166297  
 R:Kimura, M.; Sumizawa, T.; Funatsu, G.  
 Biosci. Biotechnol. Biochem. 57, 166-169, 1993  
 A:Title: The complete amino acid sequences of the B-chains of Abirin-a and Abirin-b, to  
 A:Reference number: JCI1399; MUID:93169023; PMID:7763422  
 A:Accession: JCI1399  
 A:Molecule type: protein  
 A:Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'W', 427, 'D', 429-  
 A:Experimental source: seed  
 C:Superfamily: ricin; RNA N-glycosidase homology  
 C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin  
 F:1-250/Product: abrin-b chain A #status predicted <ACH>  
 F:7-245/Domain: RNA N-glycosidase homology <RNG>  
 F:260-527/Product: abrin-b chain B #status experimental <BCH>  
 F:282-324,325-365,368-406,413-448,452-491,494-527/Region: 40-residue repeats  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted  
 F:74,113,194,195/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:110,360,400/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:163,166/Active site: Glu, Arg #status predicted  
 F:246-268,285-304,328-345,416-429,455-472/Disulfide bonds: #status predicted  
 F:287,311/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 F:499,520/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 43.1%; Score 1314; DB 2; Length 527;  
 Best Local Similarity 50.1%; Pred. No. 3.5e-92;  
 Matches 270; Conservative 77; Mismatches 178; Indels 14; Gaps 7;

40 OPIINFTAGATVQSYNTNFRAYRGRITGADVREHPIVLPNVRGPIINORFTVLSN 99  
 1 QDQVKTFTGATGTSQSYKQFTEALRELRGG--LIHIDPVLDPDTTQDENRRTVLSN 58  
 100 HAELSVTLADVTNAYVYGRAGNSAYFFH--PDNEDAEITHLFTDYNRRTFAFGNYD 159  
 59 SDTESIEAGIDVSNAYVYVAGTOSYFLRDASSASD-----YLFPTG--DOHSLPYGYGDL 114  
 162 EQLAGNLRNIEELGNPLEEASALYYSTGQPLTLARSPFICIQMSEARROYIG 221  
 117 ERNAGHQRQIPGLQALHTGIS--FFSSGGNDNEKATLTVIOMVAEARFYISN 173  
 222 EMRTIRIRNRSAPDSVTITLNSMGRSLTAIOESNOGAFASPIQLORNGSKFSYDVS 281  
 174 RYVSIOTGTAFQPDAMISLENNMNLNGVQESYQDFPNQVTLTLNRNEVYIDSLIS 233  
 282 -LIIPIIALMYKCAPPSQSLLRPVYVFNADYCDP--EPIVRIYGRNGLCYVDRD 339  
 234 HPTVAVALALMLFVCNPNPANNOSPLIRISVE--KSKICSSREPEYRIGGRGMCYVDYD 291  
 340 GREHNGNATQIMPCCKNTDANOLMTLKRNTIRSNKCLTTYGSGVYVMYIDCQAT 399  
 292 NGHNGNRITIMKCKRLBNOLMTLKSQKTIIRSNKCLTTYGAGSTVMTYIDCQATA 351  
 400 DATRMOIWDNGTITINRSSLVLAATSGNGTTLTYOTNTYAVSOGMLPTNTPOPVY 459  
 352 EATYWEIWDNGTITINRSSLVLAATSGNGTTLTYOTNTYAVSOGMLPTNTPOPVY 411  
 460 GLYGLCLQANSGQVWIEDCSSEKAEQOMALYADGSTRPOQRNDCITSDSNIRETVYK 519  
 412 GYSGLQOMAGSVMWMAADDNKKRQOMALYTDGSTRSVOANNCTLSDHKGSGTILLM 471  
 520 SCGPASSGGRMMFKNDGTITLNLXSGVLVDYRASDPISLKOIILYPLHGDNPQIMLP 576  
 472 GCGNGMASGRWYFKNDGTITLNLXSGVLVDYRASDPISLKOIILYPLHGDNPQIMLP 528





Db 82 TISVAIDVTNVMYIMGRAGDTSYF--NEASATEAKYVKDMARKVTLPSYGNVERLQ 138  
 QY 163 QLAGNLRNIEIENGPLEEALISALYYSTGCTQPLTARSFIIICIMISEARFOYIEGE 222  
 Db 139 TAAKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEARVKFIEQO 193  
 QY 223 MRRIRYRRSAPDPSPVITLSENGRSLTAIQ--ESNOGAFASPIQLORRNGSKFSV--Y 278  
 Db 194 IGRKV--DKTFPLSAIISLENSWSALSQIOIASTNNGFESPVLINQONRVITTNV 251  
 QY 279 DVSILPITIALMYR 293  
 Db 252 DAGVTSNIALLLNR 266

## RESULT 13

JU0393  
 karasurin - Mongolian snake-gourd  
 C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)  
 C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 12-Apr-1995  
 C:Accession: JU0393; PS0163  
 R:Toyokawa, S.; Takeda, T.; Kato, Y.; Wakabayashi, K.; Ogihara, Y.  
 Chem. Pharm. Bull. 39, 1244-1249, 1991  
 A:Title: The complete amino acid sequence of an abortifacient protein, karasurin.  
 A:Reference number: JU0393; MUID:92005921; PMID:1914000  
 A:Accession: JU0393  
 A:Molecule type: protein  
 A:Residues: 1-247 <TOY>  
 A>Note: a sequence which lacks Ala-247 is also shown in this publication  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 C:Keywords: abortifacient  
 F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 13.6%; Score 414; DB 2; Length 247;  
 Best Local Similarity 37.3%; Pred. No. 4,3e-24;  
 Matches 95; Conservative 61; Mismatches 81; Indels 18; Gaps 8;  
 QY 44 INFTTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNRYGLPINQRFIVELSNHAEL 103  
 Db 2 VSFRLSGATSSSYGVFISNLRKALPYERKL-YDIPLL--RSTLPQSQRVALIHLTNVADE 58  
 QY 104 SVTLALDVTNNAVYVGRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRYPFAFGNDRL 162  
 Db 59 TISVAIDVTNVMYIMGRAGDTSYF--NEASATEAKYVKDMARKVTLPSYGNVERLQ 115  
 QY 163 QLAGNLRNIEIENGPLEEALISALYYSTGCTQPLTARSFIIICIMISEARFOYIEGE 222  
 Db 116 TAAKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEARVKFIEQO 170  
 QY 223 MRRIRYRRSAPDPSPVITLSENGRSLTAIQ--ESNOGAFASPIQLORRNGSKFSV--Y 278  
 Db 171 IGRKV--DKTFPLSAIISLENSWSALSQIOIASTNNGFESPVLINQONRVITTNV 228  
 QY 279 DVSILPITIALMYR 293  
 Db 229 DAGVTSNIALLLNR 243

## RESULT 14

JC5606  
 karasurin C - Trichosanthes kirilowii var. japonica  
 M:Contains: karasurin A  
 C:Species: Trichosanthes kirilowii var. japonica  
 C>Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 19-Jul-2002  
 C:Accession: JC5606; JC5033  
 R:Mizukami, H.; Iida, K.; Kondo, T.; Ogihara, Y.  
 Biol. Pharm. Bull. 20, 711-713, 1997  
 A:Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating prote  
 A:Reference number: JC5606; MUID:97356562; PMID:9212998  
 A:Accession: JC5606  
 A:Molecule type: DNA  
 A:Residues: 1-289 <MIZ>  
 A:Cross-references: DDBJ:AB000666; NID:g2329830; PIDN:BAA21786.1; PID:g2329831

R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.  
 Biol. Pharm. Bull. 19, 1485-1489, 1996  
 A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and  
 A:Reference number: JC5032; MUID:97108848; PMID:8951169  
 A:Accession: JC5033  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 22-270 <KON>  
 C:Comment: This protein belongs to type I ribosomal-inactivating proteins which catal  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 F:22-270/Product: karasurin A #status predicted <MAC>  
 F:24-270/Product: karasurin A #status predicted <MAC>  
 F:27-266/Domain: rRNA N-glycosidase homology <RNG>

Query Match 13.6%; Score 414; DB 2; Length 289;  
 Best Local Similarity 37.3%; Pred. No. 5,5e-24;  
 Matches 95; Conservative 61; Mismatches 81; Indels 18; Gaps 8;  
 QY 44 INFTTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNRYGLPINQRFIVELSNHAEL 103  
 Db 25 VSFRLSGATSSSYGVFISNLRKALPYERKL-YDIPLL--RSTLPQSQRVALIHLTNVADE 81  
 QY 104 SVTLALDVTNNAVYVGRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRYPFAFGNDRL 162  
 Db 82 TISVAIDVTNVMYIMGRAGDTSYF--NEASATEAKYVKDMARKVTLPSYGNVERLQ 138  
 QY 163 QLAGNLRNIEIENGPLEEALISALYYSTGCTQPLTARSFIIICIMISEARFOYIEGE 222  
 Db 139 TAAKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEARVKFIEQO 193  
 QY 223 MRRIRYRRSAPDPSPVITLSENGRSLTAIQ--ESNOGAFASPIQLORRNGSKFSV--Y 278  
 Db 194 IGRKV--DKTFPLSAIISLENSWSALSQIOIASTNNGFESPVLINQONRVITTNV 251  
 QY 279 DVSILPITIALMYR 293  
 Db 252 DAGVTSNIALLLNR 266

## RESULT 15

JC5032  
 karasurin-B - Trichosanthes kirilowii var. japonica  
 C:Species: Trichosanthes kirilowii var. japonica  
 C>Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 23-May-1997  
 C:Accession: JC5032  
 R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.  
 Biol. Pharm. Bull. 19, 1485-1489, 1996  
 A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and  
 A:Reference number: JC5032; MUID:97108848; PMID:8951169  
 A:Accession: JC5032  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-247 <KON>  
 C:Comment: This protein belongs to type I ribosomal-inactivating proteins which catal  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 13.5%; Score 413; DB 2; Length 247;  
 Best Local Similarity 37.3%; Pred. No. 5,2e-24;  
 Matches 95; Conservative 63; Mismatches 79; Indels 18; Gaps 8;  
 QY 44 INFTTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNRYGLPINQRFIVELSNHAEL 103  
 Db 2 VSFRLSGATSSSYGVFISNLRKALPYERKL-YDIPLL--RSTLPQSQRVALIHLTNVADE 58  
 QY 104 SVTLALDVTNNAVYVGRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRYPFAFGNDRL 162  
 Db 59 TISVAIDVTNVMYIMGRAGDTSYF--NEASATEAKYVKDMARKVTLPSYGNVERLQ 115  
 QY 163 QLAGNLRNIEIENGPLEEALISALYYSTGCTQPLTARSFIIICIMISEARFOYIEGE 222  
 Db 116 TAAKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEARVKFIEQO 170

Tue Sep 16 12:27:46 2003

us-10-083-336a-1.rpr

Page 9

[illegible]

Search completed: September 16, 2003, 11:50:32  
Job time : 39.3567 secs

...



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:30:32; Search time 20.1928 Seconds

(without alignments)  
1341.437 Million cell updates/sec

Title: US-10-083-336a-1

Perfect score: 3051

Sequence: 1 MKPGGNTIVIMYAVATWIC.....KQILLYPLHGDPNQIWLEPLF 576

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## SUMMARIES

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	3051	100.0	576	1 RICI_RICCO	P02879 ricinus com
2	2668.5	87.5	564	1 AGGL_RICCO	P06750 ricinus com
3	1334.5	43.7	562	1 ABRC_ABRPR	P28590 abrus prece
4	1314.5	43.1	528	1 ABRA_ABRPR	P11140 abrus prece
5	1314	43.1	527	1 ABRB_ABRPR	O06077 abrus prece
6	989.5	32.4	563	1 NIGB_SAMNI	P33183 sambucus ni
7	907.5	29.7	264	1 MLB_VYSAL	P81830 viscum albu
8	428	14.0	282	1 RIP2_BRYDI	P81830 viscum albu
9	421	13.8	289	1 RIP1_TRIKI	P09989 trichosanthe
10	414	13.6	289	1 RIP1_TRIKI	P24478 trichosanthe
11	397	13.0	254	1 MIA_VYSAL	P81446 viscum albu
12	393	12.9	290	1 RIP1_BRYDI	P33185 bryonia dio
13	386	12.7	316	1 RIPG_GELMU	P33186 gellonium mu
14	375.5	12.3	286	1 RIP1_CUCFI	O9f94 cucumis fig
15	369	12.1	250	1 RIPB_LUCFY	P22851 luffa cylind
16	354.5	11.6	286	1 RIP1_LUCFY	P16094 momordica c
17	354.5	11.6	294	1 RIP1_TRIAN	P56626 trichosanthe
18	352.5	11.6	277	1 RIP1_MOMCH	O00465 luffa cylind
19	348	11.6	286	1 RIP2_MOMBA	P29339 momordica b
20	337.5	11.1	294	1 RIP1_PHYAM	O03464 phytoacta b
21	312	10.2	313	1 RIP1_PHYAM	P10297 phytoacta b
22	295.5	9.7	261	1 RIP1_PHYAM	P23339 phytoacta b
23	221.5	7.3	278	1 RIP1_PHYAM	P23339 phytoacta b
24	175.5	5.8	310	1 RIP2_MIRJA	P23339 phytoacta b
25	166	5.4	293	1 RIP0_DIACA	P23339 phytoacta b
26	164	5.4	292	1 RIP2_DIACA	P23339 phytoacta b
27	163	5.3	299	1 RIP2_SAPOF	P20656 saponaria o
28	160.5	5.3	253	1 RIP5_SAPOF	P20656 saponaria o
29	157	5.1	477	1 XYNA_SPTLI	O41389 saponaria o
30	156.5	5.1	253	1 RIP1_SAPOF	P26514 streptomyc
31	146	4.8	253	1 RIP1_SAPOF	O41391 saponaria o
32	144.5	4.7	475	1 RIP2_HORVU	P04399 hordeum vul
33	140.5	4.6	560	1 ABFB_STRCO	O54161 streptomyc
				1 JI60_HORVU	O00531 hordeum vul

34	139.5	4.6	475	1 ABFB_STRLI	P96463 streptomyc
35	139	4.6	280	1 RIP1_HORVU	P22244 hordeum vul
36	138.5	4.5	319	1 SLTA_BP933	P09385 bacterioph
37	134.5	4.4	236	1 RIP3_SAPOF	P27560 saponaria o
38	133	4.4	548	1 E13B_AKTSF	O59146 arthroacte
39	133	4.4	548	1 E13B_AKTSF	P22222 oerskovia x
40	127	4.2	545	1 AGAL_ASPNG	P25891 aspergillus
41	116	3.8	300	1 RIP3_MAIZE	P25891 zea mays (m
42	115.5	3.8	315	1 SLTA_BP930	P10149 bacterioph
43	115	3.8	301	1 RIPX_MAIZE	P28522 zea mays (m
44	114.5	3.8	315	1 SLTA_BP919	P08026 bacterioph
45	114.5	3.8	525	1 SPL_NARRA	O05308 rhabdacter

## ALIGNMENTS

## RESULT 1

RICI\_RICCO STANDARD; PRT; 576 AA.

AC P02879; P02880;

DT 21-JUL-1986 (Rel. 01, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Ricin precursor [contains: Ricin A chain (RNA N-glycosidase)

DE (EC 3.2.2.22); Ricin B chain).

OS Ricinus communis (Castor bean).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.

OX NCBI\_TaxID=3988;

RE MEDLINE=6067214; PubMed=2999712;

RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,

RT Weaver R.F.;

RT "Genomic cloning and characterization of a ricin gene from Ricinus

RT communis.";

RL Nucleic Acids Res. 13:8019-8033(1985).

RE MEDLINE=92163016; PubMed=1371405;

RA Tregear J.W., Roberts L.M.;

RT "The lectin gene family of Ricinus communis: cloning of a functional

RL ricin gene and three lectin pseudogenes.";

RL Plant Mol. Biol. 18:515-525(1992).

RE MEDLINE=85179479; PubMed=3838723;

RA Lamb A., Roberts L.M., Lord J.M.;

RT "Nucleotide sequence of cloned cDNA coding for preproricin.";

RL Eur. J. Biochem. 148:265-270(1985).

RE MEDLINE=4211267-1274(1978).

RA Yoshitake S., Funatsu G., Funatsu M.;

RT "Isolation and sequences of peptic peptides, and the complete

RL sequence of the chain of ricin-D.";

RL Agric. Biol. Chem. 42:1167-1274(1978).

RE MEDLINE=315-576.

RA Funatsu G., Kimura M., Funatsu M.;

RT "Primary structure of Ala chain of ricin D.";

RL Agric. Biol. Chem. 43:2221-2224(1979).

RE MEDLINE=90344223; PubMed=1368517;

RA Kimura Y., Kusoku H., Tada M., Takagi S., Funatsu G.;

RT "Structural analyses of sugar chains from ricin A-chain variant.";

RL Agric. Biol. Chem. 54:157-162(1990).

RE MEDLINE=21480122; PubMed=11595634;

RA Olshes S., Kozlov J.V.;

RT "Ricin.";  
 RL Toxicon 39:1723-1728(2001).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=87165983; PubMed=3558397;  
 RA Montfort W., Villafraña J.E., Monzingo A.F., Ernst S.R., Katzin B.,  
 RA Rutenber E., Xiong N.H., Hamlin R., Robertus J.D.;  
 RL "The three-dimensional structure of ricin at 2.8 Å.";  
 RT J. Biol. Chem. 262:5398-5403(1987).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=91352004; PubMed=1881881;  
 RA Katzin B.J., Collins E.J., Robertus J.D.;  
 RL "Structure of ricin A-chain at 2.5 Å.";  
 RT Proteins 10:251-259(1991).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.  
 RX MEDLINE=91352005; PubMed=1881882;  
 RA Rutenber E., Robertus J.D.;  
 RL "Structure of ricin B-chain at 2.5-Å resolution.";  
 RT Proteins 10:260-269(1991).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=95082010; PubMed=7990130;  
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,  
 RA Pauplt R.A.;  
 RL "X-ray structure of recombinant ricin A-chain at 1.8-Å resolution.";  
 RT J. Mol. Biol. 244:410-422(1994).  
 RN [12]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.  
 RX MEDLINE=96374222; PubMed=8780513;  
 RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,  
 RA Molina-Svint M.C., Robertus J.D.;  
 RL "Structure and activity of an active site substitution of ricin A chain.";  
 RT Biochemistry 35:11098-11103(1996).  
 RN [13]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=97240820; PubMed=9086280;  
 RA Yan X., Hollis T., Svint M., Day P., Monzingo A.F., Milne G.W.,  
 RA Robertus J.D.;  
 RL "Structure-based identification of a ricin inhibitor.";  
 RT J. Mol. Biol. 266:1043-1049(1997).  
 RN [14]  
 RP MUTAGENESIS.  
 RX MEDLINE=93165632; PubMed=1287657;  
 RA Kin Y., Robertus J.D.;  
 RL "Analysis of several key active site residues of ricin A chain by mutagenesis and X-ray crystallography.";  
 RT Protein Eng. 5:775-779(1992).  
 CC -1- FUNCTION: Ricin is highly toxic to animal cells and to a less extent to plant cells. The A chain is responsible for inhibiting protein synthesis through the catalytic inactivation of 60S ribosomal subunits. It acts as a glycosidase that removes a specific adenine residue from an exposed loop of 28S ribosomal RNA. As this loop is involved in the binding of elongation factors, the modified ribosomes are unable to support protein synthesis. The A chain can inactivate a few thousand ribosomes per minute, thus inactivating them faster than the cell can make new ones. A single A-chain molecule can therefore kill an animal cell. The B chain binds to cell receptors and facilitates the entry into the cell of the A chain; B chains are also responsible for cell agglutination (lectin activity). It binds to beta-D-galactopyranoside moieties.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: Disulfide-linked dimer of A and B chains.  
 CC -1- DOMAIN: The B chain is composed of two domains, each domain consists of 3 homologous subdomains (alpha, beta, gamma).  
 CC -1- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.

CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -1- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).  
 CC -1- DATABASE: NAME-Protein Spotlight;  
 CC NOTE=Issue 31 of February 2003;  
 CC WWW="http://www.expasy.org/spotlight/articles/sp11031.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: X03179; CAA26939.1; -;  
 DR EMBL: X52908; CAA37095.1; -;  
 DR EMBL: X02388; CAA26230.1; -;  
 DR EMBL: A12892; CAA01058.1; -;  
 DR PIR: A24041; RLCSO.  
 DR PDB: 2AAT; 31-JAN-94.  
 DR PDB: 1ARG; 31-JAN-94.  
 DR PDB: 1EMP; 31-OCT-93.  
 DR PDB: 1IFS; 14-JAN-98.  
 DR PDB: 1IFU; 14-JAN-98.  
 DR PDB: 1IRC; 31-OCT-93.  
 DR PDB: 1ORS; 16-JUN-97.  
 DR PDB: 1OBT; 16-JUN-97.  
 DR PDB: 1BR5; 02-SEP-98.  
 DR PDB: 1BR6; 02-SEP-98.  
 DR PDB: 1IL3; 16-JAN-02.  
 DR PDB: 1IL4; 16-JAN-02.  
 DR PDB: 1IL9; 16-JAN-02.  
 DR GlycoSuiteDB: P02879; -;  
 DR InterPro: IPR00772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; Signal; 3D-structure.  
 FT SIGNAL 1 35  
 FT CHAIN 36 302  
 FT PEPTIDE 303 314  
 FT CHAIN 315 576  
 FT DOMAIN 321 448  
 FT DOMAIN 451 575  
 FT REPEAT 331 373  
 FT REPEAT 374 414  
 FT REPEAT 417 449  
 FT REPEAT 462 497  
 FT REPEAT 501 540  
 FT REPEAT 543 570  
 FT ACT\_SITE 212 212  
 FT DISULFID 294 318  
 FT DISULFID 334 353  
 FT DISULFID 377 394  
 FT DISULFID 465 478  
 FT DISULFID 504 521  
 FT CARBOHYD 45 45  
 FT CARBOHYD 271 271  
 FT CARBOHYD 409 409  
 FT CARBOHYD 449 449  
 FT CONFLICT 76 76  
 FT CONFLICT 551 551  
 FT STRAND 43 47  
 FT TURN 49 50  
 N-LINKED (GLCNAC. . .).  
 /FTID-CAR.000080.  
 N-LINKED (GLCNAC. . .) (IN MINOR FORM).  
 /FTID-CAR.000081.  
 N-LINKED (GLCNAC. . .).  
 N-LINKED (GLCNAC. . .).  
 E -> D (IN REF. 3).  
 A -> R (IN REF. 3).  
 INTERCHAIN.

Query Match 100.0%; Score 3051; DB 1; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 3e-233;  
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKPGNTIYIMWAAVATMLCGSTSGMSFTLEDNNIFPKOYPIINFTAGATVOSTYTFI 60  
 1 MKPGNTIYIMWAAVATMLCGSTSGMSFTLEDNNIFPKOYPIINFTAGATVOSTYTFI 60  
 61 RAVRRLTGADVREIPLVLRNRCPLPINOPIVLELSNAHLSVTLALDVTNAYVGYR 120  
 61 RAVRRLTGADVREIPLVLRNRCPLPINOPIVLELSNAHLSVTLALDVTNAYVGYR 120  
 61 RAVRRLTGADVREIPLVLRNRCPLPINOPIVLELSNAHLSVTLALDVTNAYVGYR 120  
 121 AGNSAFEPHPNOEDAEALITLFTDVQNRRTFAFGANDRLEOLAGNIRENIEGNGPLE 180  
 121 AGNSAFEPHPNOEDAEALITLFTDVQNRRTFAFGANDRLEOLAGNIRENIEGNGPLE 180  
 181 EALSALYVSTGGTQPLPLANSFTICIQMISEAARFOYIEGEMETRIYRNRASAPDSVI 240  
 181 EALSALYVSTGGTQPLPLANSFTICIQMISEAARFOYIEGEMETRIYRNRASAPDSVI 240  
 241 TLENSWGRSTAIOESNOGAFASPIOLORNGSKFSYVDVSLIPITALMYRCAPPSS 300  
 241 TLENSWGRSTAIOESNOGAFASPIOLORNGSKFSYVDVSLIPITALMYRCAPPSS 300  
 301 QPSLLIRPVNPNADVCMDEPIVIRIVGRNGLCVDVDRGFRHNGNAIQMPCKSNPDAN 360  
 301 QPSLLIRPVNPNADVCMDEPIVIRIVGRNGLCVDVDRGFRHNGNAIQMPCKSNPDAN 360  
 361 QMLTKLRNTRRSNGKCLTGYSPGYVYVYDCNTAATATATROQINDNGTIIPRSSLY 420  
 361 QMLTKLRNTRRSNGKCLTGYSPGYVYVYDCNTAATATATROQINDNGTIIPRSSLY 420  
 421 LAATSGNSGTLTVOYTNIVASOGMLPTNTOPTVITVIGLGLCLANSQOWIEPCSS 480  
 421 LAATSGNSGTLTVOYTNIVASOGMLPTNTOPTVITVIGLGLCLANSQOWIEPCSS 480  
 481 EKAQOQALYADGSIROQNRDNCLESDSNIREFVVKILSCGPASSGQRMFKNDGTILN 540  
 481 EKAQOQALYADGSIROQNRDNCLESDSNIREFVVKILSCGPASSGQRMFKNDGTILN 540  
 541 LYSGLVLDVRAVSDPSLKOIILYPLHGDPNOIWLPLF 576  
 541 LYSGLVLDVRAVSDPSLKOIILYPLHGDPNOIWLPLF 576

RESULT 2  
 AGGL\_RICCO STANDARD; PRT; 564 AA.  
 AC P06750:  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Agglutinin precursor (RCA) [contains: Agglutinin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Agglutinin B chain].  
 OS Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae; eucotsid I; Malpighiales; Euphorbiaceae; Ricinus.  
 NCBI\_TaxID=3988;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86059449; PubMed=2999130;  
 RA Roberts L.M., Lamb F.T., Pappin D.J.C., Lord J.M.;  
 RT "The primary sequence of Ricinus communis agglutinin. Comparison with ricin.";  
 RL J. Biol. Chem. 260:15682-15686(1985).  
 RP SEQUENCE OF 303-564.  
 RC TISSUE=Seed.  
 RA Ataki T., Yoshioaka Y., Funatsu G.;  
 RT "The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds.";  
 RT

RL Biochim. Biophys. Acta 872:277-285(1986).  
 RN [3]  
 RP SEQUENCE OF 303-337.  
 RX MEDLINE=80178723; PubMed=6768555;  
 RA Lin T.T.-S., Li S.S.-L.;  
 RT "Purification and physicochemical properties of ricins and agglutinins from Ricinus communis.";  
 RL Eur. J. Biochem. 105:453-459(1980).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M12089; AAA33869.1; -;  
 DR EMBL: S40368; AAB22584.1; -;  
 DR PIR: A24261; RICSAG.  
 DR HSSP: P02879; 1BR6.  
 DR GYCOSULEDB: P06750; -;  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHICARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B-LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA-RICIN; 1.  
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; Signal.  
 FT CHAIN 1 24  
 FT PROPEP 25 290  
 FT CHAIN 291 302  
 FT CHAIN 303 364  
 FT DOMAIN 309 436  
 FT DOMAIN 439 563  
 FT REPEAT 319 361  
 FT REPEAT 362 402  
 FT REPEAT 405 437  
 FT REPEAT 450 485  
 FT REPEAT 489 528  
 FT REPEAT 531 558  
 FT ACT\_SITE 200 200  
 FT DISULFID 282 306  
 FT DISULFID 322 341  
 FT DISULFID 365 382  
 FT DISULFID 453 466  
 FT CARBOHYD 34 34  
 FT CARBOHYD 259 259  
 FT CARBOHYD 397 397  
 FT CARBOHYD 437 437  
 FT CONFLICT 331 331  
 FT CONFLICT 362 362  
 FT CONFLICT 374 374  
 FT CONFLICT 404 404  
 FT CONFLICT 552 552  
 SQ SEQUENCE 564 AA; 62851 MW; D45F2A72F609759 CRC64;

Query Match 87.5%; Score 2668.5; DB 1; Length 564;  
 Best Local Similarity 89.6%; Pred. No. 4.8e-203;  
 Matches 506; Conservative 26; Mismatches 32; Indels 1; Gaps 1;

12 MVAATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTAGATVOSTYTFIRAVRRLTGGA 71  
 1 MVAATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTAGATVOSTYTFIRAVRRLTGGA 71

QY	72	DVRHHEPLVLPNRVNGLPINORFLIVELSHNAELSTLADVTNMYVVGVRAGNSAYCFEHD	133
Db	61	DVRHHEPLVLPNRVNGLPISORFLIVELSHNAELSTLADVTNMYVVGCRGNSAIFPFHD	120
QY	132	NOEDAEATLHLETTDVQNRVTFAFGGNRYRLQOLAGNIRENLETGNGPLEPAISALYYST	191
QY	192	GCTOLTFLARSTIICIMISEAARQYIEGMRIRIRYNRRSADPSVITLNSWGLST	251
Db	180	CGTQIPTFLARSPVCIOMISEAARQYIEGMRIRIRYNRRSADPSVITLNSWGLST	239
QY	252	AIQESNOGAFAPIQLORRNGSKSEVYVSTLIPILAMVRCAPPSQSFLLIRPVY	311
Db	240	AIQESNOGAFAPIQLORRNGSKSEVYVSTLIPILAMVRCAPPSQSFLLIRPVY	299
QY	312	NFNADVCMPPEPVIIVIRVGRNGLCYDVRGRFRHNGNAIOLMPCSKSNTDANOLMLKRDNT	371
Db	300	NFNADVCMPPEPVIIVIRVGRNGLCYDVGEFFEDGNPILMPCSKSNTDANOLMLKRDNT	359
QY	372	RSNGKCLTTYGSPGVYVMIYDCNTAATDATRMOJWDNGTILNFRSSLVLAATSGNGTT	431
Db	360	RSNGKCLTTYGSPGVYVMIYDCNTAATDATRMOJWDNGTILNFRSGVLAATSGNGTK	419
QY	432	LFTQNTIVAVSGWMLPTNNTQPFVTVTIGVGLCLQANSGQVWLEDCSSEKAEQOMALYA	491
Db	420	LFTQNTIVAVSGWMLPTNNTQPFVTVTIGVGLCMCLQANSGVWLEDCSEKAEQOMALYA	479
QY	492	DGSIIRPOQNRDNCILSDSNIRETVYKILSCGPASSGQRMWRKNDGTILNLVSGVLVDYRA	551
Db	480	DGSIIRPOQNRDNCILTDANIKGTVYKILSCGPASSGQRMWRKNDGTILNLVSGVLVDYRR	539
QY	552	SDPSLKQIILYPLHGDPNQLPLF 576	
Db	540	SDPSLKQIILYPLHFGNQLPLF 564	

RESULT 3

ABRC.ABRR

ID	ABRC.ABRR	STANDARD;	PRT;	562 AA.
AC	P28590;			
DT	01-DEC-1992 (Rel. 24, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Abtrin-c precursor [Contains: Abtrin-c A chain (tRNA N-glycosidase)			
DE	(EC 3.2.2.22), Abtrin-c B chain].			
OS	Abrus precatorius (Indian licorice) (Crab's eye).			
OC	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;			
OC	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;			
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.			
OX	NCBI_TaxID=3816;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Leaf;			
RX	MEDLINE=91266957; PubMed=2050149;			
RA	Wood K.A., Lord J.M., Mawrzycazk E.J., Platak M.;			
RT	"Preproabrin: genomic cloning, characterisation and the expression of			
RL	the A-chain in Escherichia coli.";			
RL	Eur. J. Biochem. 198;723-732(1991).			
CC	- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN			
CC	SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL			
CC	SUBUNTS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE			
CC	B CHAIN IS A GLYCOSIDE-SPECIFIC LECTIN THAT FACILITATES THE			
CC	BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.			
CC	- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one			
CC	specific adenosine on the 28S rRNA.			
CC	- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.			
CC	- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN			
CC	CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).			
CC	- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-			
CC	INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.			
CC	- SIMILARITY: Contains 2 ricin B-type lectin domains.			

	CC	-----	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as the content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch</a> ).
	CC	-----	
	DR	EMBL; X55667; CAA39202.1; -	
	DR	PIR; S16022; S16022.	
	DR	HSSP; P11140; IABR.	
	DR	InterPro; IPRO00772; Ricin_B_lectin.	
	DR	InterPro; IPRO01574; RIP.	
	DR	Pfam; PF00652; Ricin_B_lectin; 6.	
	DR	Pfam; PF00161; RIP; 1.	
	DR	PRINTS; PRO0396; SHIGARICIN.	
	DR	SMART; SM00458; RICIN; 2.	
	DR	PROSITE; PS50231; RICIN_B_LECTIN; 2.	
	DR	PROSITE; PS00273; SHIGA_RICIN; 1.	
	KW	Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;	
	KW	Glycoprotein; Lectin; Signal; Pyridoxone carboxylic acid.	
	FT	SIGNAL	1 34 BY SIMILARITY.
	FT	CHAIN	35 285 ABRIN C A CHAIN (BY SIMILARITY).
	FT	PEPTIDE	286 295 LINKER PEPTIDE (BY SIMILARITY).
	FT	CHAIN	286 562 ABRIN B C CHAIN (BY SIMILARITY).
	FT	DOMAIN	307 434 RICIN B-TYPE LECTIN 1.
	FT	DOMAIN	437 561 RICIN B-TYPE LECTIN 2.
	FT	REPEAT	317 359 1-ALPHA.
	FT	REPEAT	360 400 1-BETA.
	FT	REPEAT	403 435 1-GAMMA.
	FT	REPEAT	448 483 2-ALPHA.
	FT	REPEAT	487 526 2-BETA.
	FT	REPEAT	529 562 2-GAMMA.
	FT	ACET_SITE	198 198 BY SIMILARITY.
	FT	DISULFID	281 303 INTERCHAIN (BY SIMILARITY).
	FT	DISULFID	320 339 BY SIMILARITY.
	FT	DISULFID	363 380 BY SIMILARITY.
	FT	DISULFID	451 464 BY SIMILARITY.
	FT	DISULFID	480 507 BY SIMILARITY.
	FT	MOD_RES	35 PYRIDOXONE CARBOXYLIC ACID (BY SIMILARITY).
	FT	CARBOHYD	234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
	FT	CARBOHYD	435 435 N-LINKED (GLCNAC. .) (POTENTIAL).
	SO	SEQUENCE	562 AA; 62817 MW; 1FD0ABC7D7BA6278 CRC64;
		Query Match	43.7%; Score 1334.5; DB 1; Length 562;
		Best Local Similarity	48.9%; Pred. No. 1,le-97;
		Matches 276; Conservative	80; Mismatches 191; Indels 17; Gaps 8
QY	19	LCFGSTSGMSPFLDENNIIP----	-KOYLPIINTTAGAVQSNTFRRAVRGRLTGADV R 74
DB	10	LCLAMTCNSALRCARARYPPVATNODQIKRTTGAGANSOSKQTEALRLRGTFG--LI	67
QY	75	HEIPLVLPNRVGILPNORFTLVLSNHAEISVTALAIDVTANAYVYGRAGSAIEFFIPDNDE	134
DB	68	HDIPVLDPDTVEEERKRYITVELSENSESIENVGIDVTNAVVAVARAGSQSYEL--RDA	124
QY	135	DKEATHLEPTDVONRTFAFGNGNYRDELQAGNLBNIELNGSPLEAFISAIVYYSTGGT	194
DB	125	PASASTYLETPGTQ-RISLRFDSYGDLRWMAQTREELISGLQALTHAIS---FLRSAGS	180
QY	195	OPLTAATSFLLTIICOMISEAARQYITEGEMTRIRYNRRNSADPPSVYTILENSGRUSTAIO	254
DB	181	NDEEKARTLIIVIOMASEAARYIISNRGVISRCTGARQPDPAMSLTENMDNISGVO	240
QY	255	ESNKGAFASPDIOLQRNSKFYSVDVS-ILIPILIMVYRKAPPSQSSLIRPVPMF	313
DB	241	OSVQDTFPNNVILLISTINROPVVVSDSHPTVALALMLEVCNPNNANOSPPLIRSYVE--	298
QY	314	NAVVCNDP--EPYRIYGRNGLCDVADGDFRHNGNALIQMPCKSNTDANOQLTKRDNTIR	372

```

Db      299  ESKICSSREPTVRIGSGDGCVDYDDGNGNRIIAMCKDRLEENQMTLTKSDKTI R 358
Qy      373  SNGKCLTITGYSPGVYWIYDCNTAATDATRMQIMDNGTILNPRSLVATSGNSGTTL 432
Db      359  SNGKLTTEGAPGVYWIYICTSAVAETATWEIMDGTILNPRSLVATSGNSGTTL 418
Qy      433  TVQNTIYAVSOGWMLPTNNTOPFTVTYGLVGLCLQANSQGVYIEDCSSEKAQOMALYAD 492
Db      419  TVQNTIYAVSOGWMLPTNNTOPFTVTYGLVGLCLQANSQGVYIEDCSSEKAQOMALYAD 478
Qy      493  GSIRPOQRNDCLTSDSNIRRTVYKILSCGRASSGQRMENRNDGTILNLVSGLVLYVRAS 552
Db      479  GSISVQNTNCLTSKHKQSPYVLMACSNMASQRMFLFKNDGSIVNLHDDVMYDKRS 538
Qy      553  DPSLKQILYPLHGDPNQIMLPLF 576
Db      539  DPSLKELIHPHKGKPNQIMLPLF 562

RESULT 4
ABRA_ABRPR STANDARD: PRT: 528 AA.
AC P1140: P28589:
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Abirin-a precursor (Contains: Abirin-a A chain (rRNA N-glycosidase)
DE (EC 3.2.2.22); Abirin-a B chain).
OS Abirus preclatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotyledons 1; Fabales; Fabaceae; Papilionoideae; Abraceae; Abrus.
OX NCBI_Taxid=3816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93132798; PubMed=8421313;
RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
RT "Primary structure of three distinct isoabirins determined by cDNA
RT sequencing. Conservation and significance.";
RL J. Mol. Biol. 229:263-267(1993).
RN [2]
RP SEQUENCE OF 1-251.
RC TISSUE=Seed;
RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
RT "The complete amino acid sequence of the A-chain of abirin-a, a toxic
RT protein from the seeds of Abrus precatorius.";
RL Agric. Biol. Chem. 52:1095-1097(1988).
RN [3]
RP SEQUENCE OF 1-251 FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=91201329; PubMed=2016300;
RA Evensen G., Mathiesen A., Sundan A.;
RT "Direct molecular cloning and expression of two distinct abirin
RT A-chains.";
RL J. Biol. Chem. 266:6848-6852(1991).
RN [4]
RP SEQUENCE OF 262-528.
RX MEDLINE=92371656; PubMed=1505674;
RA Chen Y.-L., Chow L.-P., Tsungita A., Lin J.-Y.;
RT "The complete primary structure of abirin-a B chain.";
RL FEBS Lett. 309:115-118(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).
RX MEDLINE=95333188; PubMed=7608980;
RA Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
RT "Crystal structure of abirin-a at 2.14 A.";
RL J. Mol. Biol. 250:354-367(1995).
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC ABRIN-A IS MORE TOXIC THAN RICIN.
CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT

```

```

CC PRECEDES ENDOCYTOSIS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.
CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M98344; AAA3624.1; ALT_INIT.
DR EMBL: X54872; -: NOT_ANNOTATED_CDS.
DR PIR: S32429; TRLSA.
DR PDB: IABR; 07-FEB-95.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN.2.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein; Lectin; 3D-structure; Pyrrolidone carboxylic acid.
FT CHAIN 1 251
FT PEPTIDE 252 261
FT CHAIN 262 528
FT DOMAIN 273 400
FT DOMAIN 403 527
FT REPEAT 283 325
FT REPEAT 326 366
FT REPEAT 369 401
FT REPEAT 414 449
FT REPEAT 453 492
FT REPEAT 495 528
FT ACT_SITE 164 164
FT DISULFID 246 269
FT DISULFID 285 305
FT DISULFID 329 346
FT DISULFID 417 430
FT DISULFID 456 473
FT MOD_RES 1 1
FT CARBOHYD 361 361
FT CONFLICT 202 202
FT CONFLICT 298 298
FT CONFLICT 427 427
FT CONFLICT 467 467
FT CONFLICT 483 483
FT STRAND 5 8
FT TURN 10 11
FT HELIX 14 28
FT STRAND 32 33
FT TURN 34 35
FT STRAND 36 38
FT TURN 42 43
FT HELIX 47 49
FT STRAND 51 57
FT STRAND 63 69
FT TURN 70 72
FT STRAND 75 79
FT STRAND 83 86
FT TURN 88 89
FT TURN 92 93
FT HELIX 94 97

```

FT TURN 100 101  
 FT STRAND 103 106  
 FT TURN 113 114  
 FT HELIX 115 119  
 FT TURN 124 126  
 FT STRAND 129 129  
 FT HELIX 131 142  
 FT TURN 143 144  
 FT HELIX 148 167  
 FT STRAND 168 168  
 FT HELIX 169 180  
 FT TURN 181 182  
 FT STRAND 185 185  
 FT HELIX 189 196  
 FT TURN 197 197  
 FT HELIX 198 207  
 FT STRAND 212 220  
 FT TURN 222 223  
 FT STRAND 226 231  
 FT TURN 232 233  
 FT HELIX 235 239  
 FT STRAND 240 240  
 FT STRAND 243 243  
 FT STRAND 248 248  
 FT STRAND 268 268  
 FT STRAND 276 277  
 FT STRAND 279 280  
 FT HELIX 282 284  
 FT STRAND 286 289  
 FT HELIX 290 292  
 FT TURN 296 297  
 FT STRAND 299 303  
 FT HELIX 311 313  
 FT STRAND 315 317  
 FT TURN 319 320  
 FT STRAND 322 325  
 FT TURN 326 327  
 FT STRAND 328 332  
 FT TURN 337 338  
 FT STRAND 340 344  
 FT TURN 346 348  
 FT HELIX 351 353  
 FT STRAND 355 355  
 FT STRAND 357 358  
 FT TURN 360 361  
 FT STRAND 364 366  
 FT TURN 367 370  
 FT STRAND 371 374  
 FT TURN 380 381  
 FT STRAND 383 383  
 FT STRAND 385 387  
 FT HELIX 393 395  
 FT STRAND 398 399  
 FT STRAND 406 408  
 FT STRAND 410 411  
 FT HELIX 413 415  
 FT STRAND 417 421  
 FT TURN 422 423  
 FT STRAND 424 428  
 FT TURN 432 433  
 FT HELIX 435 437  
 FT STRAND 439 441  
 FT TURN 443 444  
 FT STRAND 447 449  
 FT TURN 450 451  
 FT STRAND 452 459  
 FT TURN 464 465  
 FT STRAND 467 472  
 FT TURN 474 475  
 FT HELIX 478 480  
 FT STRAND 483 484  
 FT TURN 486 487  
 FT STRAND 490 492

FT TURN 493 496  
 FT STRAND 497 501  
 Query Match 43.18; Score 1314.5; DB 1; Length 528;  
 Best Local Similarity 49.7%; Pred. No. 3.8e-96;  
 Matches 267; Conservative 78; Mismatches 175; Indels 17; Gaps 8;  
 QY 44 INFETGATVQSYNTFLRAVGRLTGADVRRHPIPLPVRVGLPIRQRFILVELSNHAEI 103  
 DB 5 IKFSTEGATQSOSTKQTEALRELRGG--LIHDIPLPDPPTIIOENRRIITVELSDPE 62  
 QY 104 SVTLALDVTNAVYVYVAGNSAYFFH--PDNDADAIHLEFTDVQNRRTFAFGNYDL 161  
 DB 63 SIEVGIDVTNAVYVAVRAGTQSYFLDAPASAD----YLFQGT-DQHSLEFGYGYDL 116  
 QY 162 EQLAGNIRENIEIGNGPLEAIGALYYISGTQTLTARSFICIQMISEARQYIEG 221  
 DB 117 ERMAHOSKQOIPILGLALTHGIS--FFRSGNDNEKARTLIVLIQVAAEARFRYISN 173  
 QY 222 EMFRIRYNRSAPDPSVITPLENSMGRLESTAIOESNOGAFASP1QIQRNGSKFSYDVS 281  
 DB 174 RVKYSIDGTGAFPQPDAMISLENMNDLSRGVESYQDPFPNQVTLTNIRNEPVIYDLS 233  
 QY 282 -ILPIIALMVYRCAPPSSQFSLLRPVYPNENADYCDP-EPYIRYRNGLCVYVD 339  
 DB 234 HPTVAVLALMLFCNPENANOSPLIRISIVE--KSRICSRPEPTVIRIGRGMCVYVD 291  
 QY 340 GRFHNGNALIOLMPCKSSTNDANOLWTLKRDNTIRSNCKLTPTYGSPGVYVMIYDCFTAT 399  
 DB 292 NGYHNGRRIITMKCKDBLENOJLTKSDTIRSNCKLTPTYGSPGVYVMIYDCFTAT 351  
 QY 400 DATEWQIMDNGTIIINPRSSIVLAATGNSGTLTVOTNIYAVSQWLPTNTOPEVTTIV 459  
 DB 352 EATYWEIWDNGTIIINPRSSIVLAATGNSGTLTVOTNEIYAVSQWLPTNTOPEVTTIV 411  
 QY 460 GLYGCLQANSQGVWIDCSSEKAEQOMAIYADGSTRPOQNRNCLTSNIRETVKTL 519  
 DB 412 GYSDLCMQAQSNVWMAADCDSNKRQOMALYTGDSIRSVQNTNNCLTSKDHQGSFTILM 471  
 QY 520 SCGPASSGQRMWFKNDGTIINLYSGVLVDPRASPSLKOIILYPLHGDPOIWLPLF 576  
 DB 472 GCSNGMASQRMVFRKNDGSIISLYDDVMYDKGSDPSLKOIILYPLHGDPOIWLPLF 528  
 RESULT 5  
 ABRB\_ABRPR STANDARD; PRT; 527 AA.  
 ID ABRB\_ABRPR  
 AC 006077; p81374;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Abrin-b precursor (Contains: Abrin-b A chain (rRNA N-glycosidase))  
 DE (EC 3.2.2.22); Abrin-b B chain).  
 OS Abrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Adreae; Abrus.  
 OC NCBI\_TaxID=3816;  
 RX MEDLINE=93132798; PubMed=8421313;  
 RA Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;  
 RT "Primary structure of three distinct isoabins determined by cDNA  
 J. Mol. Biol. 229:263-267(1993).  
 RN [2]  
 RP SEQUENCE OF 260-527.  
 RC TISSUE=Seed;  
 RX MEDLINE=93169023; PubMed=7763422;  
 RA Kimura M., Sumizawa T., Funatsu G.;  
 RT "The complete amino acid sequences of the B-chains of abrin-a and  
 abrin-b, toxic proteins from the seeds of Abrus precatorius.";  
 RL Biosci. Biotechnol. Biochem. 57:166-169(1993).

CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
CC SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
CC ABRIN-A IS MORE TOXIC THAN RICIN  
CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
CC PRECEDES ENDOCYTOSIS.  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
CC specific adenosine on the 28S RNA.  
CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
CC INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.  
CC -1- SIMILARITY: Contains 2 Ricin B-type lectin domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL, M98345; AAA32625.1; -.  
DR PIR, S32430; S32430.  
DR HSSP, P11440; IABR.  
DR InterPro: IPR000772; Ricin\_B\_lectin.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR SMART: SM00458; RICIN; 2.  
DR PROSITE: PS0231; RICIN\_B\_LECTIN; 2.  
DR PROSITE: PS00275; SHIGA-RICIN; 1.  
KW Plant defense: Hydrolyase; Protein synthesis inhibitor; Toxin; Repeat;  
KW Glycoprotein; Lectin; Pyrrolidone carboxylic acid.  
KW CHAIN 1 250 ABRIN-B A CHAIN.  
FT PEPTIDE 251 260  
FT CHAIN 261 527 ABRIN-B B CHAIN.  
FT DOMAIN 272 399 RICIN B-TYPE LECTIN 1.  
FT DOMAIN 402 526 RICIN B-TYPE LECTIN 2.  
FT REPEAT 282 324 1-ALPHA.  
FT REPEAT 325 365 1-BETA.  
FT REPEAT 368 400 1-GAMMA.  
FT REPEAT 413 448 2-ALPHA.  
FT REPEAT 452 491 2-BETA.  
FT REPEAT 494 527 2-GAMMA.  
FT ACT SITE 163 163  
FT DISULFID 246 268  
FT DISULFID 285 304  
FT DISULFID 328 345  
FT DISULFID 416 429  
FT DISULFID 455 472  
FT MOD\_RES 1 1  
FT CARBOHYD 110 110  
FT CARBOHYD 360 360  
FT CARBOHYD 400 400  
FT CARBOHYD 282 282  
FT CONFLICT 291 291  
FT CONFLICT 350 351  
FT CONFLICT 378 378  
FT CONFLICT 426 426  
FT CONFLICT 428 428  
FT CONFLICT 431 431  
FT CONFLICT 484 484  
FT CONFLICT 491 491  
FT CONFLICT 493 493  
FT CONFLICT 502 502  
FT CONFLICT 509 509  
FT CONFLICT 513 513  
FT CONFLICT 516 516  
N- LINKED (GLCNAC. . .) (POTENTIAL).  
N- LINKED (GLCNAC. . .) (POTENTIAL).  
N- LINKED (GLCNAC. . .) (POTENTIAL).  
N- > D (IN REF. 2).  
D- > N (IN REF. 2).  
AE- > PO (IN REF. 2).  
S- > N (IN REF. 2).  
L- > M (IN REF. 2).  
Y- > D (IN REF. 2).  
N- > S (IN REF. 2).  
R- > K (IN REF. 2).  
N- > S (IN REF. 2).  
H- > Y (IN REF. 2).  
R- > G (IN REF. 2).  
E- > Q (IN REF. 2).  
H- > W (IN REF. 2).  
H- > T (IN REF. 2).

SG SEQUENCE 527 AA; 59114 MW; 3253AE490CE949A CRC64;  
Query Match 43.18; Score 1314; DB 1; Length 527;  
Best Local Similarity 50.14; Pred. No. 4.1e-96;  
Matches 270; Conservative 77; Mismatches 178; Indels 14; Gaps 7;  
QY 40 QYPIINFTTGAIVOSTNFRIRAVRGRLTGGAVRHEIPVLPKRVGLPINORFIVLELN 99  
D 1 QDOVTKFTTCAGTSQSOKFIEALRORLTGCG--LHGIPVLPPTLOERNRIVSELN 58  
QY 100 HAEISVTLADVTNAVYVGYRAGNSAVFFHPDQDEAEATHTLFTVDQNRYPFAGND 159  
D 59 SDRESIAGIDVSNAYVAVRAONRGFL--RDAPTSARYLFTGTQ-QYSLRFGSTI 114  
QY 160 RLEQLAGNLRNLELNGPLLEAISALYYSTGCTQLPLARSEFIICQIMISEARFOYI 219  
D 115 DLERLAKOTRQOPIGLIGLALRHAISFL---QSGTDDQETARLLIVIQMASAPARYRI 170  
QY 220 EGEKRTIRYRNRSPDPDYITLENMGRLSTIOESNCGAFASPIQLORNKSKSYVD 279  
D 171 SYRGVSIRNTVFPQDAAMISLENMNDLSGVQOQVDDTPFNATVLTQIMASAPARYRI 230  
QY 280 VS-TLPIITIALVYRCAPPPSSQSLIRPVYFNVADYCMPD-EPIVIRVGRNGLCVY 337  
D 231 LTHQSAVALALMLFCNPNPANGSPILLISIVE--KSKICSSRYEPTVIRIGRNGCVY 288  
QY 338 RDGRFHNGNATIQLPCKSNTDANQTLTKRDNTRNSGKCLTGYSPGVYMIYDCNTA 397  
D 289 YDGYHNGNRIIAMKCKDRLEENQWLTKSDKTRNSGKCLTEGAPGNYMIYDCNTA 348  
QY 398 ATDATRMOIWDNGTIIINPSSILVLAATSGNSGTTLVONTIYAVSGCWLEPTNTOPEVYT 457  
D 349 VAAETWEIWDNGTIIINPKSALVLSSESSMGSTLTVOYNEIIMRGWGTGNTSEFVTS 408  
QY 458 IYGLVGLQIANSQGVWIEDCSSEKAEQOMALYADGSIIRPQNRDNCILTSNIRETVYK 517  
D 409 ISGYSGLCQAOQSNVLAIVCDNNKKEQOMALYTDGSIISYQNTNCLTSKHKQGSPIV 468  
QY 518 IISGCRASSGQGRMFNNDGTLNLYSGVLVDYRASPSSKQIILPLHDDPQWLPLEF 576  
D 469 IMAQSGMASQRMFLPNDGSIYNLHDDVMDVYKRSPLKEITILHYHGRPNQWLTLEF 527  
RESULT 6  
NIGB\_SAMNI STANDARD; PRT; 563 AA.  
ID NIGB\_SAMNI  
AC P33183; P33184; P93542;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Nigirin b precursor (Agglutinin V) (SNMV) [contains: Nigirin b A chain  
DE (rRNA N-glycosidase) (EC 3.2.2.22); Nigirin b B chain].  
OS Sambucus nigra (European elder).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
OX NCBI\_TaxId=4202;  
RP SEQUENCE FROM N.A.  
RC TISSUE-Bark;  
RC MEDLINE=96215449; PubMed=8647092;  
RA Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;  
RT "Characterization and molecular cloning of Sambucus nigra agglutinin V  
RT (nigirin b), a GalNAc-specific type-2 ribosome-inactivating protein  
RT from the bark of elderberry (Sambucus nigra).";  
RL Eur. J. Biochem. 237:505-513(1996).  
RN [2]  
RP SEQUENCE OF 26-49 AND 298-321.  
RC TISSUE-Bark;  
RC MEDLINE=94003077; PubMed=8400135;  
RA Gibbs T., Citores L., Ferreras J.M., Rojo M.A., Iglesias R.,  
RA Munoz R., Arias P.J., Calonge M., Garcia J.R., Mendez E.;  
RT "Isolation and partial characterization of nigirin b, a non-toxic







DR PROSITE: PS50231; RICIN\_B-LECTIN; 2.  
 KW Plant defense; Glycoprotein; Lectin; Repeat; 3D-structure.  
 FT DOMAIN 8 135 RICIN B-TYPE LECTIN 1.  
 FT DOMAIN 139 263 RICIN B-TYPE LECTIN 2.  
 FT BINDING 23 25 GALACTOSE.  
 FT BINDING 236 238 GALACTOSE.  
 FT DISULFID 5 5 INTERCHAIN.  
 FT DISULFID 64 81 BY SIMILARITY.  
 FT DISULFID 152 165 BY SIMILARITY.  
 FT DISULFID 192 209 BY SIMILARITY.  
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 18 18 N -> S.  
 FT VARIANT 56 56 G -> N.  
 FT VARIANT 157 157 G -> O.  
 FT VARIANT 195 195 S -> Y.  
 FT VARIANT 224 224 G -> Y.  
 FT VARIANT 231 231 N -> S OR T.  
 FT VARIANT 231 233 NGL -> KGP.  
 FT VARIANT 232 235 GLAM -> SLMV.  
 FT SEQUENCE 264 AA; 28981 MM; 7D0DC326CCE6F5A4 CRC64;

Query Match 29.7%; Score 907.5; DB 1; Length 264;  
 Best Local Similarity 63.8%; Pred. No. 2e-64;  
 Matches 166; Conservative 33; Mismatches 60; Indels 1; Gaps 1;

QY 318 CMDEPTVRIYRGNGLCVDRDGRFHNGNAIOLMPCKSNFNDQMTLRKNDTRIRNGKC 377  
 Db 5 CSASEPTVRIYRGNGMVDVDRDDFDHNGNOIOLMPKSNNDPMQMTIRKNDTRIRNGSC 64  
 QY 378 LTTGYSPGYVYVYDCNTAT 437  
 Db 65 LTTGYTAGYVYVYDCNTAT 124  
 QY 438 IYVSGMGLPTNNTOPFTTIVGGLCLQANGSGVWIEEC-SSEKAEQOMALYAGSIR 496  
 Db 125 DYTLGGWGLAGNTPAREVTIYGFRLCLMSNGSVWETCDSSQKNQGWALYAGSIR 184  
 QY 497 PQONRNDCLTSDSNIRETVKILSCGPASSGQRMFKNDGTIINTLXSGILDVRAADPSL 556  
 Db 185 PKONODCLTSGRDSVTVINIVSCGASGSRMVFTEBNAIINTLNGLAMVDVAQNPRL 244  
 QY 557 KOIILVPLHGDPNOMLPLF 576  
 Db 245 RIITITPATGKPNQMWLPVF 264

## RESULT 8

RIP2\_BRVDI STANDARD; PRT; 282 AA.  
 AC P8184; 09S8J0; 40, Created)  
 DT 16-OCT-2001 (Rel. 40, last sequence update)  
 DT 28-FEB-2003 (Rel. 41, last annotation update)  
 DE Ribosome-inactivating protein bryodin II precursor (rRNA N-glycosidase) (EC 3.2.2.22) (BD2).  
 OS Bryonia dioica (Red Bryonia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.  
 NCBI\_TaxID=1652;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stegall C.B., Gawlak S.L., Marguardt H.;  
 RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant Bryonia dioica."  
 RL Patent number US5597569, 28-JAN-1997.  
 RN [2]  
 RP SEQUENCE OF 22-42.  
 RC TRISSE-Root; PubMed=7849072;  
 RX MEDLINE=95151812; PubMed=7849072;  
 RA Stegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,

RA Marguardt H.;  
 RT "Characterization of ribosome-inactivating proteins isolated from Bryonia dioica and their utility as carcinoma-reactive immunoconjugates."  
 RT Bioconjug. Chem. 5:423-429(1994).  
 CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 1 RIP SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to ilcense@sib-sib.ch) or send an email to ilcense@sib-sib.ch).  
 CC EMBL: 134238; -; NOT\_ANNOTATED\_CDS.  
 DR HSP; P09989; 1MRJ.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
 KW Multigene family; Glycoprotein; Signal.  
 FT CHAIN 1 21  
 FT ACT\_SITE 183 183 RIBOSOME-INACTIVATING PROTEIN BRYODIN II.  
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 282 AA; 30754 MM; C52BE2P6A873769C CRC64;

Query Match 14.0%; Score 428; DB 1; Length 282;  
 Best Local Similarity 43.3%; Pred. No. 1.6e-26;  
 Matches 110; Conservative 41; Mismatches 85; Indels 18; Gaps 10;

QY 44 INFTAGATVOSYTFIRAVRGRLTGTGADVREIPVLN-RVGLPINDRFLVELSNIAE 102  
 Db 24 INFSLGATGATYKFFIRLRLKLTGVPRYIDIVLNNAAGL--ARFOLVLTITNNG 80  
 QY 103 ISVTLADVTNAYVYVYGRAGNSAVFPHPDNEADAEITLFTDQONRTFFAGGNDYRL 162  
 Db 81 ESVTVALDVYVYVYVAVRAGNTAYFL--ADASTANNVLEAGI-NHVRLEYGNDYRL 136  
 QY 163 QLAGNL-RENTPLGNGPLBEAISALYYSTGCTQLPTLARSFIIQIMISAPRYYTEG 221  
 Db 137 TAAGRISRENTPLGSEISSAIGNMFRNP-GTSVP--RAFIYIQTVSEARFKYI-- 190  
 QY 222 EMKTRIRYRRSAPPSPVITLNSMGRSTAIQ--ESNQGAFASPIOLORNGSKFSYVD 279  
 Db 191 EDRVSENVGTAKFPAPLFLQNMAGSLSEQIQAQTRGGEFARVELKTVSNTPFTVTN 250  
 QY 280 V-SILIPITALMV 292  
 Db 251 VNSPVVKGIALLLY 264

## RESULT 9

RIP2\_TRIKI STANDARD; PRT; 289 AA.  
 AC P09989;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-NOV-1990 (Rel. 16, last sequence update)  
 DT 15-SEP-2003 (Rel. 42, last annotation update)  
 DE Ribosome-inactivating protein alpha-trichosanthin precursor (rRNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS).  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 NCBI\_TaxID=3677;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Maximowicz;  
 RX MEDLINE=91153657; PubMed=1999291;  
 RA Shaw P.C., Yang M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;  
 RT "Cloning of trichosanthin cDNA and its expression in *Escherichia*  
 RT coli.";  
 RL Gene 97:267-272(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Maximowicz; TISSUE=Leaf;  
 RX MEDLINE=90256790; PubMed=2341400;  
 RA Chow T., Feldman R.A., Lovett M., Piatk M.;  
 RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a  
 RT type I ribosome-inactivating protein.";  
 RL J. Biol. Chem. 265:8670-8674(1990).  
 RN [3]  
 RP SEQUENCE OF 24-270.  
 RC STRAIN=Maximowicz; TISSUE=Tubercous root;  
 RX MEDLINE=90256789; PubMed=2341399;  
 RA Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,  
 RA Wu P., Hwang K., Piatk M.;  
 RT "Primary amino acid sequence of alpha-trichosanthin and molecular  
 RT models for abrin A-chain and alpha-trichosanthin.";  
 RL J. Biol. Chem. 265:8665-8669(1990).  
 RN [4]  
 RP SEQUENCE OF 24-270.  
 RC TISSUE=Tubercous root;  
 RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,  
 RA Tian G.Y., Ni C.Z.;  
 RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and  
 RT application.";  
 RL Pure Appl. Chem. 58:789-798(1986).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).  
 RX MEDLINE=94344957; PubMed=8066085;  
 RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;  
 RT "Structure of trichosanthin at 1.88-A resolution.";  
 RL Proteins 19:4-13(1994).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
 RX MEDLINE=95344383; PubMed=7619070;  
 RA Huang O., Liu S., Tang Y., Jin S., Wang Y.;  
 RT "Studies on crystal structures, active-centre geometry and  
 RT depurinating mechanism of two ribosome-inactivating proteins.";  
 RL Biochem. J. 309:285-296(1995).  
 CC -I- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS  
 CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.  
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M34858; AAA34207.1; -;  
 DR EMBL: J05434; AAA34206.1; -;  
 DR PIR: J05566; RLTXP.  
 DR PDB: 1MRJ; 07-FEB-95.  
 DR PDB: 1MRK; 07-FEB-95.  
 DR PDB: 1TCS; 10-JUL-95.  
 DR PDB: 1J4G; 28-JAN-03.  
 DR PDB: 1NLI; 21-JAN-03.  
 DR PDB: 1OD2; 24-APR-00.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.

DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN: 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 270  
 FT PROPEP 271 289  
 FT ACT\_SITE 183 183  
 FT CONFLICT 57 60  
 FT CONFLICT 82 84  
 FT CONFLICT 87 87  
 FT CONFLICT 92 92  
 FT CONFLICT 143 144  
 FT CONFLICT 196 196  
 FT CONFLICT 215 216  
 FT CONFLICT 231 231  
 FT CONFLICT 234 234  
 FT CONFLICT 246 266  
 FT CONFLICT 247 247  
 FT STRAND 25 28  
 FT TURN 30 31  
 FT HELIX 34 46  
 FT TURN 47 47  
 FT STRAND 50 54  
 FT TURN 55 56  
 FT STRAND 57 60  
 FT HELIX 66 69  
 FT STRAND 70 76  
 FT TURN 78 79  
 FT STRAND 82 88  
 FT TURN 89 92  
 FT STRAND 93 99  
 FT TURN 100 101  
 FT STRAND 102 105  
 FT HELIX 109 114  
 FT TURN 115 117  
 FT STRAND 120 121  
 FT STRAND 124 127  
 FT HELIX 134 141  
 FT TURN 142 142  
 FT HELIX 145 147  
 FT STRAND 150 150  
 FT HELIX 152 163  
 FT TURN 164 165  
 FT HELIX 167 180  
 FT TURN 181 181  
 FT HELIX 182 186  
 FT STRAND 187 187  
 FT HELIX 188 195  
 FT TURN 196 196  
 FT STRAND 202 202  
 FT HELIX 206 226  
 FT TURN 227 230  
 FT STRAND 231 239  
 FT TURN 241 242  
 FT STRAND 245 250  
 FT TURN 251 252  
 FT HELIX 254 258  
 FT TURN 259 259  
 FT STRAND 260 260  
 FT STRAND 263 263  
 FT TURN 266 268  
 SQ SEQUENCE 289 AA; 31676 MW; 5CE09BB6305758B9 CRC64;  
 Query Match 13.8%; Score 421; DB 1; Length 289;  
 Best Local Similarity 37.3%; Pred. No. 6e-26;  
 Matches 95; Conservative 60; Mismatches 82; Indels 18; Gaps 8;  
 QY 44 INFETGATGVOSTYTFNFIKAVRGRLLTGADVRHEIPVLPKRVGLPIKQRFIVLELSNHAL 103  
 Db 25 VSFRLSGATSSSYGVFISNLKRLPNERKL-YDIPLL--RSSLPGSQRYALIHILNAYDE 81

FT	CHAIN	22	270	KARASURIN-C.
FT	CHAIN	22	270	KARASURIN-C.
FT	PROPEP	271	289	REMOVED IN MAJURE FORM.
FT	ACT_SITE	183	183	BY SIMILARITY.
SO	SEQUENCE	289 AA:	31704 MM;	883D3E3242887B26 CRC64;
Query Match				
Best Local Similarity		13.6%;	Score 414;	DB 1: Length 289;
Matches		95;	Conservative 61;	Mismatches 81; Indels 18; Gaps 8;
QY	44	INFTAGATVOSTYNTFIRAVGRLLTGADVHEIFVLNPNRGLPNQRFIVETSNAHEL	103	
DB	25	VSFRISGATSSSYGVFISNLKRALPYEKL-YDIPDL--RSTLPQSYXALHILTNVDE	81	
QY	104	SVTLALDVTNAYVGYRAGNSATFFHPNOEDA-EATHLFTDVONRYTFAFGNYDRLE	162	
DB	82	TISVAIDTNYVMKRYAGDTSYFF--NEASATEAAKYVFKDAKRYKLTLPYSGNERYLQ	138	
QY	163	QLAGNLRENIENGLNGLPEEATLSALYYSTGGTOLPTLARSPFIIOMISEARQYIEGE	222	
DB	139	IAAGIRIRNITPLGLPALDASATTLTFYIAN-----SAASALMWLIQSTSEARRYKLEQO	193	
QY	223	MRTIRIKRRSAPDPSPVITLLENSWGRSLTAIQ--ESNQGAPASPIOLRRNGSKFSV--Y	278	
DB	194	IGKRY--DKTFLPSTAIISLSEWSALSKQIQIASTNGGQETPEVVLINQNGVITITNV	251	
QY	279	DVSIILPIIATMAYR	293	
DB	252	DAGVVTSMIALLLNR	266	
RESULT 11				
MLA_VISAL	ID	MLA_VISAL	STANDARD:	PRT: 254 AA.
AC	P81446.			
DT	15-DEC-1998	(Rel. 37, Created)		
DT	15-DEC-1998	(Rel. 37, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Beta-galactoside specific lectin I A chain (MLA) (ML-I A) (rRNA N-			
DE	glycosidase) (EC 3.2.2.22).			
OS	Viscum album (European mistletoe).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Santalales; Santalaceae; Viscum.			
OX	NCBI_TaxID=3972;			
RN	[1]			
RP	SEQUENCE.			
RC	SPRAIN-Subsp. album.			
RX	MEDLINE=97134581; Pubmed=8980141;			
RA	Soler M.H., Stoeva S., Schwaborn C., Wilhelm S., Stiefel T.,			
RA	Voelter W.;			
RT	*Complete amino acid sequence of the A chain of mistletoe lectin I.*;			
RL	FEBS Lett. 399:153-157(1996).			
CC	-1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN			
CC	SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL			
CC	SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE			
CC	B CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY			
CC	INTO THE CELL OF THE A CHAIN. B CHAINS ARE ALSO RESPONSIBLE FOR			
CC	CELL AGGLUTINATION (LECTIN ACTIVITY).			
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one			
CC	specific adenosine on the 28S rRNA.			
CC	-1- SUBUNIT: Disulfide-linked dimer of A and B chains.			
CC	-1- PHARMACOLOGICAL: Due to its immunomodulative effects it is being			
CC	studied in clinical trials in cancer patients as it may slow the			
CC	growth of cancer cells and be an effective treatment for solid			
CC	tumors.			
CC	-1- MISCELLANEOUS: TWO ISOPFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA			
CC	AND NON-GLYCOSYLATED FORM MLA'.			
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.			
CC	TYPE 2 RIP SUPERFAMILY.			
CC	PIR: PD0018; PD0018.			
CC	HSSP: P1140; 1ABR.			
CC	Interpro: IPR00157; IPR			

DR Pfam; PF00161; RIP, 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; FALSE\_NEG.  
 KW Plant defense; Hydrolase; Toxin; Repeat; Glycoprotein; Lectin.  
 FT ACT\_SITE 165 165  
 FT CARBOHYD 112 112  
 FT VARIANT 15 15  
 FT VARIANT 15 15  
 FT VARIANT 66 66  
 FT VARIANT 112 112  
 FT VARIANT 116 116  
 FT VARIANT 133 133  
 FT VARIANT 140 140  
 FT VARIANT 144 144  
 FT VARIANT 151 151  
 FT VARIANT 179 179  
 FT VARIANT 184 184  
 FT VARIANT 190 190  
 FT VARIANT 218 218  
 FT VARIANT 223 224  
 FT VARIANT 231 231  
 FT VARIANT 235 235  
 SQ SEQUENCE 254 AA; 28478 MW; 53BAF98D3E0FEE67 CRC64;  
 Query March 13.0%; Score 397; DB 1; Length 254;  
 Best Local Similarity 40.8%; Pred. No. 3.9e-24;  
 Matches 104; Conservative 45; Mismatches 94; Indels 12; Gaps 8;  
 QY 48 TAGATVQSYTNFIRAVRGRLTTGADVREHPIPLNRYGLPIN--QRPITVELSNHAEISV 105  
 DB 9 THQTGEYEFITLLRDYVSSG--FSNEIPL--RSTIPVSDAQRFVLELITNOGDSV 66  
 QY 106 TLADVTNAYVYGRAGNSAEFFIPDNOEDAEATHFTFDQNKNTYFAFGNRYRLPOLA 165  
 DB 67 TLAIDVNAVYVAOAGQSYFLR-DAPRGAE--THFTGT--TSSLPFNSYIDLRRYA 122  
 QY 166 GNLRNIEINGNGPLEEASIALYVYSTGTQPLTLARSEFIICQIMISEARFOYIEGEMRT 225  
 DB 123 GH-RQPIPLGIDQLQSVTALRF--PGSGRTQARSILLIQMISEARFNPLIMRYRQ 178  
 QY 226 RIRYRRASAPDPVSITLLENSGRSLTAIOESNQGAFSPIDQRRNSKFSYVYSLIP 285  
 DB 179 YNSGASPLPVMYMETLSWGOOSTOVHSTDGVFNPIRLAIPGFTVLINVRDYIA 238  
 QY 286 ITALMVYRCAPPPSS 300  
 DB 239 SLAIMLFVCGERPPS 253  
 RESULT 12  
 RIPL\_BRYDI STANDARD; PRT; 290 AA.  
 AC P33185; O9S819;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein bryodin I precursor (rRNA N-glycosidase)  
 DE (EC 3.2.2.22) (BDI).  
 OS Bryonia dioica (Red bryony).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC euroids I; Cucurbitales; Cucurbitaceae; Bryonia.  
 OX NCBI\_TaxID=3652;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
 RC TISSUE=leaf;  
 RX MEDLINE=97228081; PubMed=9115985;  
 RA Siegall C.B.;  
 RA Siegall C.B.;  
 RT "Molecular, biological, and preliminary structural analysis of  
 RT recombinant bryodin I, a ribosome-inactivating protein from the plant  
 RT Bryonia dioica".  
 RL Biochemistry 36:3095-3103(1997).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA Siegall C.B.;  
 RT "Cloning and expression of a gene encoding bryodin I from Bryonia  
 RT dioica".  
 RL Patent number US5541110, 30-JUL-1996.  
 RN [3]  
 RP SEQUENCE OF 24-66.  
 RC TISSUE=Seed;  
 RX MEDLINE=89326691; PubMed=2753596;  
 RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,  
 RA Lappi D.;  
 RT "N-terminal sequence of some ribosome-inactivating proteins";  
 RT Int. J. Pept. Protein Res. 33:263-267(1989).  
 RN [4]  
 RP SEQUENCE OF 24-43.  
 RC TISSUE=Root;  
 RX MEDLINE=95151812; PubMed=7849072;  
 RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,  
 RA Marguardt H.;  
 RT "Characterization of ribosome-inactivating proteins isolated from  
 RT Bryonia dioica and their utility as carcinoma-reactive  
 RT immunokonjugates";  
 RL Bioconjug. Chem. 5:423-429(1994).  
 CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
 CC PROTEIN SYNTHESIS IN ANIMAL CELLS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO  
 CC PRODUCE A SHORTER PROTEIN.  
 CC -1- BIOCHEMISTRY: Especially useful as immunotoxin for  
 CC pharmacological applications as it has low toxicity in rats and  
 CC mice but is potent once inside target cells.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
 CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
 CC -----  
 DR EMBL; I24020; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; S16491; S16491.  
 DR PDB; 1BR7; 04-MAR-98.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
 KW 3D-structure; Multigene family; Glycoprotein; Signal.  
 FT SIGNAL 1 23  
 FT CHAIN 24 270  
 FT PROPEP 271 290  
 FT ACT\_SITE 183 183  
 FT ACT\_SITE 212 212  
 FT CARBOHYD 214 214  
 FT CARBOHYD 250 250  
 FT MUTAGEN 212 212  
 FT CONFLICT 61 65  
 FT STRAND 25 28  
 FT STRAND 30 31  
 FT TURN 30 31  
 FT TURN 34 34  
 FT HELIX 46 46  
 FT TURN 47 47  
 FT STRAND 50 54  
 FT TURN 55 55  
 FT STRAND 57 57  
 FT STRAND 66 69  
 FT HELIX 66 69  
 FT STRAND 70 76  
 FT TURN 78 79  
 FT STRAND 82 88  
 FT TURN 89 92  
 CC -----  
 CC RIBOSOME-INACTIVATING PROTEIN BRYODIN I.  
 CC MISSING IN MATURE PROTEIN.  
 CC BY SIMILARITY.  
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC E->K: REDUCES ACTIVITY 10-FOLD.  
 CC RSSIS -> LRHXI (IN REF. 3).  
 CC -----

FT STRAND 93 99  
 FT TURN 100 101  
 FT STRAND 102 105  
 FT HELIX 109 114  
 FT TURN 115 117  
 FT TURN 120 121  
 FT STRAND 124 127  
 FT HELIX 134 141  
 FT TURN 142 142  
 FT HELIX 145 147  
 FT STRAND 150 150  
 FT HELIX 152 163  
 FT TURN 164 165  
 FT STRAND 167 186  
 FT HELIX 187 187  
 FT STRAND 188 196  
 FT STRAND 202 202  
 FT HELIX 206 213  
 FT TURN 214 214  
 FT HELIX 215 225  
 FT TURN 226 230  
 FT STRAND 231 239  
 FT TURN 241 242  
 FT STRAND 245 250  
 FT TURN 251 252  
 FT HELIX 254 257  
 FT TURN 258 259  
 FT STRAND 260 260  
 FT STRAND 263 263  
 FT HELIX 266 268  
 SQ SEQUENCE 290 AA; 31788 MW; E966CD9C031A42DB CRC64;

Query Match 12.9%; Score 393; DB 1; Length 290;  
 Best Local Similarity 33.7%; Pred. No. 9.8e-24;  
 Matches 98; Conservative 67; Mismatches 92; Indels 34; Gaps 10;

QY 8 IVIMVAVATMLCGSTGSGSFTLEDNNIPKQPIINFTAGATVOSTNEIRAVKRL 67  
 Db 5 LVMLLLITLITFL-----KSPVTEGD-----VSFRLSGATTTSTGVFTKRLREAL 48  
 QY 68 TTGADVHEIPVLPNRYGLPTINOFILVELSNHAEISVTLALVTNAYVVGYSAGSAYF 127  
 Db 49 PYERKV-YNIPPL--RSISGSGRYTLHLTNADEITSAVDVTNYITGLADVSIF 105  
 QY 128 FHPDNDGA-EAITHLEFDVONRYTAFGNGYDLRLQALNLENREITELGNGPLEAISAL 186  
 Db 106 F--NEASATEAKAFYFKDAKKVTLPSGNYERLQTAACKIRENIPILGLALDSALITTL 162  
 QY 187 YYYSTGTQPLTARSFIICIMISEARQYIEGEMRTIRIKYRNSAPDPSTYITLNSG 246  
 Db 163 YTTAS-----SNASALLVLIQSTASARAKFIEQIGKRV--DKTFPLSLATISLENNW 215  
 QY 247 GRUATAIO--ESNOGAFASPIQLORRNGSKFSYDVS--ILPIIALMYR 293  
 Db 216 SALSKOIQAISTNNQFESFPVVLIDQNNQVSTITNASARVYSINALLNLR 266

RESULT 13  
 RIBG\_GELMU STANDARD; PRT; 316 AA.  
 ID RIBG\_GELMU STANDARD; PRT; 316 AA.  
 AC P3186;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein gelonin precursor (rRNA N-glycosidase)  
 DE (EC 3.2.2.22).  
 GN GEL.  
 OS Gelonium multiflorum (Euphorbiaceae himalayae).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Gelonium.  
 NCBI\_Taxid=3979;  
 [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=94085781; PubMed=7916721;  
 RA Nolan P.A., Garrison D.A., Beller M.;  
 RT "Cloning and expression of a gene encoding gelonin, a ribosome-  
 inactivating protein from Gelonium multiflorum.";  
 RL Gene 134:223-227(1993).  
 RN [2]  
 RP SEQUENCE OF 47-93.  
 RC TISSUE=Seed.  
 RX MEDLINE=89326691; PubMed=2753596;  
 RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,  
 RA Lappi D.;  
 RT "N-terminal sequence of some ribosome-inactivating proteins.";  
 RL Int. J. Pept. Protein Res. 33:263-267(1989).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
 RX MEDLINE=95333189; PubMed=7608981;  
 RA Hosur M.V., Nair B., Satyamurthy P., Misquith S., Surobia A.,  
 RA Kannan K.K.;  
 RT "X-ray structure of gelonin at 1.8-A resolution.";  
 RL J. Mol. Biol. 250:368-380(1993).  
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -I- SUBUNIT: Homodimer.  
 CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
 CC -----  
 DR EMBL: J12243; AAA16312.1; -  
 DR PIR: J12243; J12243.1;  
 DR HSSP: P09989; IMR7.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP.1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN.  
 DR KMW Plant defense: Protein synthesis inhibitor; Hydrolase; Toxin; Signal;  
 KMW glycoprotein.  
 FT SIGNAL 1 26  
 FT PROPEP 27 46  
 FT CHAIN 47 297  
 FT PROPEP 298 316  
 FT DISULFID 90 96  
 FT CARBOHYD 235 235  
 FT ACT\_SITE 212 212  
 FT CONFLICT 90 90  
 FT CONFLICT 93 93  
 SQ SEQUENCE 316 AA; 35418 MW; 1252F3E710901B85 CRC64;  
 Query Match 12.7%; Score 386; DB 1; Length 316;  
 Best Local Similarity 35.0%; Pred. No. 3.9e-23;  
 Matches 116; Conservative 54; Mismatches 129; Indels 32; Gaps 15;  
 QY 5 GNTIIVIM-VAVATMLC-----FGSTSG-WSTFLEDNNIPKQPI--INFTAGATVOS 55  
 Db 3 GNMKYYWIKIIVATWFCCTTIVLGSTARIFSLPNDDEBETSKTGLDVTSGATYIT 62  
 QY 56 YTNFIRAVRGRLTGADVHEIPVLPNRYGLPTINOFILVELSNHAEISVTLADVTNAY 115  
 Db 63 YNFIENLRVRLKPEGN--SHGIPILRKCDP--GKCEVLALSDNQLAEIALIDVSY 120  
 QY 116 VVGYSAGSAYFEH--PDNODAEAIHTLFTD-VONRYTAFGNGYDLRLQALN--LRE 170  
 Db 121 VVGIVQVRNRSYFEKQADP-----AAYEGLEFNITKTR--LHFGGSYSLE---GEKAYRE 170  
 QY 171 NIELGNGLEAEIALISLTYITSGTQPLTARSFIICIMISEARQYIEGEMRTIRYN 230  
 [1]

QY 44 INTTGAAYGVOSTNFRVAVRGTLTGADVRIHEHPVL--PVRVLPINORFIELNSHAE 102  
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 28 VRFSLGSSNNKSKTSKITSMRNLMPAGDI-VNIPLLVPISG---SRRTILMQLSNYEG 83  
  
QY 103 LSTLTALDPTNAVVGGRAGNSAYPEFHNDQDEAEATHTFTVQNRKYTAFCGNDRLE 162  
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 84 NITTMADVDNVNIMGLVNGTSTFE---NEIDQLASKRVFGGTSTILPEYSGNOKLO 140  
  
QY 163 QLAGNLRRENIETLNGPLPEEAISALYYVSTGGTOLPTLARSTTICIQMISFAARFYOTEGE 222

Qy 224 RTRIRYNNRRSAPPVITLEN-SWGRISTAIO--ESNOGAFASPIQLOÖRRNGSKESFYDY 280  
 || | - : : ||| : | | | : : ||| : : :  
 Db 172 IERIPKN--EVSPALSTLENFAWSLTSKOIOLOAÖTNGCAFRTPVVILDNKGÖRVETITL 229  
 || | : : ||| : | | | : : ||| : : :  
 Qy 281 SILIPi 286  
 : : : |  
 Db 230 ASKVQI 235

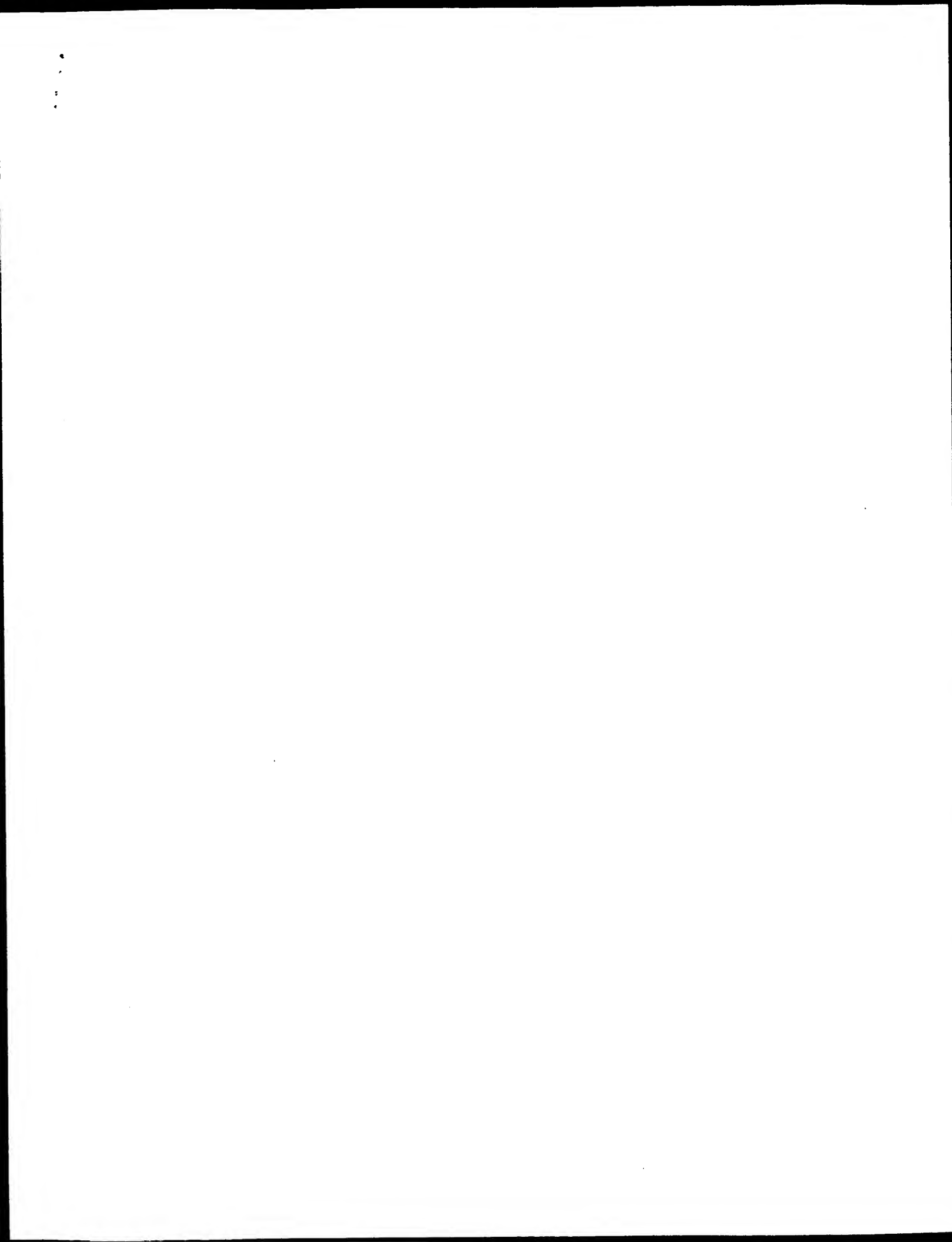
Tue, Sep 16 12:27:46 2003

us-10-083-336a-1.rsp

Page 15

Search completed: September 16, 2003, 11:46:01  
Job time : 22.1928 secs

---





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:40:08 : Search time 91.8773 Seconds  
(without alignments)  
1617.791 Million cell updates/sec

Title: US-10-083-336a-1  
Perfect score: 3051  
Sequence: 1 MKPGNTYIVMYAVATWLC.....KQILLYPLHGDPNQWLPWF 576

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SPREMBL\_23:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mmc:\*  
9: sp\_organelle:\*  
10: sp\_phase:\*  
11: sp\_plant:\*  
12: sp\_ricent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*  
16: sp\_virus:\*  
17: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2843	93.2	541	10	041174 ricinus com
2	1492	48.9	580	10	094bw3 cinnamomum
3	1487	48.7	580	10	094bw4 cinnamomum
4	1443.5	47.3	581	10	094bw5 cinnamomum
5	1441.5	47.2	549	10	094bw5 cinnamomum
6	1341.5	44.0	263	10	041143 ricinus com
7	1317.5	43.2	528	10	006076 abrus precia
8	1315	43.1	531	10	08RXH6 abrus precia
9	1286	42.2	555	10	08W243 viscum albu
10	1244	40.8	547	10	09W6F9 viscum albu
11	1109	36.3	592	10	08W2E7 abrus precia
12	1081	35.4	573	10	08W2E7 iris hollan
13	1065	34.9	564	10	08W2E8 iris hollan
14	1060.5	34.8	563	10	09AVR2 sambucus eb
15	1053	34.5	382	10	004367 sambucus ni
16	1000.5	32.8	563	10	08SA43 abrus precia
					08GT32 sambucus ni

17	989.5	32.4	563	10	094552 sambucus ni
18	983.5	32.2	565	10	004071 sambucus ni
19	976.5	32.0	566	10	004072 sambucus ni
20	971.5	31.8	570	10	022415 sambucus ni
21	947.5	31.1	570	10	041358 sambucus ni
22	942	30.9	263	10	08LKQ2 viscum albu
23	933.5	30.6	604	10	09W654 polyonatum
24	915	30.0	263	10	08LKQ1 viscum albu
25	906	29.7	603	10	09W653 polyonatum
26	820.5	26.9	569	10	P93543 sambucus ni
27	785.5	25.7	266	10	08LKQ3 viscum albu
28	641	21.0	316	10	08GTA5 sambucus ni
29	640	21.0	316	10	093WL1 sambucus ni
30	639	20.9	316	10	094554 sambucus ni
31	638	20.9	320	10	094553 sambucus ni
32	635	20.8	320	10	004366 sambucus ni
33	621.5	20.4	307	10	08GTA6 sambucus ni
34	451	14.8	252	10	038760 abrus precia
35	429	14.1	252	10	038761 abrus precia
36	423	13.9	251	10	096236 abrus precia
37	422	13.8	251	10	096237 abrus precia
38	419	13.7	289	10	094KE4 trichosanthe
39	416	13.6	251	10	096235 abrus precia
40	414	13.6	289	10	041216 trichosanthe
41	411	13.5	247	10	09LRE3 trichosanthe
42	398.5	13.1	278	10	000980 luffa cylin
43	398	13.0	270	10	08LPV7 trichosanthe
44	395	12.9	270	10	041611 trichosanthe
45	393	12.9	254	10	08LKQ6 viscum albu

## ALIGNMENTS

RESULT 1  
ID 041174  
AC 041174: PRELIMINARY; PRT; 541 AA.  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Protein A chain (EC 3.2.2.22) (rRNA N-glycosidase)  
DE (Fragment).  
OS Ricinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92338377; PubMed=1633311;  
RA Roberts L.M., Tregear J.W., Lord J.M.;  
RT Targeted cloning of ricin.;  
RL Targeted Diagon. Ther. 7:81-97(1992).  
CC -I- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RNA.  
CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
CC EMBL: S40366; AAB2582.1; -.  
DR HSSP: P02879; 1BR6.  
DR InterPro: IPR000772; Ricin\_B\_lectin.  
DR InterPro: IPR001574; RIP.  
DR InterPro: IPR001400; Somatotropin.  
DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR SMART: SM00438; RICIN; 2.  
DR PROSITE: PS50231; RICIN\_B\_LLECTIN; 2.  
DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
DR PROSITE: PS00338; SOMATOTROPIN\_2; 1.  
KW Hydrolase; Toxin.  
FT NON\_TER 1  
SQ SEQUENCE 541 AA: 60281 MW: 28782CDEF1F2E9D9 CRC64;

Query Match	93.2%	Score 2843;	DB 10;	Length 541;
Best Local Similarity	99.6%;	Pred. No. 7.8e-224;		
Matches 539; Conservative	1;	Mismatches	1;	Indels 0; Gaps 0.

QY	3	IPPKOYPIINFTTAATVOSTYNFNRARGLTGTGADVRHEITPLPVRGGLPINORILY	95
Db	1	IPPKOYPIINFTTAATVOSTYNFNRARGLTGTGADVRHEITPLPVRGGLPINORILY	60
QY	96	ELSNHAEISVTLALDVTNAYVVGYPAGNSAFYFHPDNOEDAEATLHFTVONKRYTFAFG	155
Db	61	ELSNHAEISVTLALDVTNAYVVGYPAGNSAFYFHPDNOEDAEATLHFTVONKRYTFAFG	120
QY	156	GWYDRELOAGLRENIELGNGLPEEALISALYYTSTGGTQPLTARSTICIMISAAR	215
Db	121	GWYDRELOAGLRENIELGNGLPEEALISALYYTSTGGTQPLTARSTICIMISAAR	180
QY	216	FOYIEGEMTRIRYNNRSAPDPSPYLTLENSMGRISTAIOESNOGAFASPIOLORNSKF	275
Db	181	FOYIEGEMTRIRYNNRSAPDPSPYLTLENSMGRISTAIOESNOGAFASPIOLORNSKF	240
QY	276	SVYDVSILPIIALWVRYKCAPPPSSQSFLLRPVVPNNADVCMDPEYIRIYGRNGLCY	335
Db	241	SVYDVSILPIIALWVRYKCAPPPSSQSFLLRPVVPNNADVCMDPEYIRIYGRNGLCY	300
QY	336	DVRDGFHNGNALIQIMPCKSNTDANQMLTKRDNTINSNGCLTTYGYSPQYVMIYDCN	395
Db	301	DVRDGFHNGNALIQIMPCKSNTDANQMLTKRDNTINSNGCLTTYGYSPQYVMIYDCN	360
QY	396	TATDTRMQINDNGTIIINPRSLVLAATSGNSGTTLVQYNITAYVSOQMLPTNNTOPFV	455
Db	361	TATDTRMQINDNGTIIINPRSLVLAATSGNSGTTLVQYNITAYVSOQMLPTNNTOPFV	420
QY	456	TTTVGLYGLCLQANSQVWIEDCSSEKAEDQOMALYADGSIRPOONRDNCULTSDSNIREV	515
Db	421	TTTVGLYGLCLQANSQVWIEDCSSEKAEDQOMALYADGSIRPOONRDNCULTSDSNIREV	480
QY	516	VKILSGPASSQQRWFKNDGTIILLYSGVLVDNRASPSLKOIILYPLHGDPNQWMLPL	575
Db	481	VKILSGPASSQQRWFKNDGTIILLYSGVLVDNRASPSLKOIILYPLHGDPNQWMLPL	540
QY	576	F 576	
Db	541	F 541	

RESULT 2		
094BM3		
ID	094BM3	PRELIMINARY; PRT; 580 AA.
AC	094BM3;	
DT	01-DEC-2001 (TREMBLrel. 19, Created)	
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)	
DE	Type 2 ribosome-inactivating protein cinnamominn III precursor	
DE	(EC 3.2.2.22) (rRNA N-glycosidase).	
OS	Cinnamomum camphora (Camphor tree).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliopsida; Laurales; Lauraceae; Cinnamomum.	
OX	NCBI_taxid=13429;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Yang O., Gong Z.Z., Liu W.Y.;	
RT	"Molecular cloning of three type 2 RIP (ribosome-inactivating protein)	
RT	genes encoding cinnamominn proteins and study of their expression	
RT	patterns.";	
RL	Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.	
CC	-1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE	
CC	SPECIFIC ADENOSINE ON THE 28S RRNA.	
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.	
DR	EMBL; AY039803; AAK82460.1; "	
DR	InterPro; IPR000772; Ricin_B_lectin.	
DR	InterPro; IPR001574; RIP	
DR	Pfam; PF00652; Ricin_B_lectin; 6.	
DR	Pfam; PF00161; RIP; 1.	

DR	PRINTS: P000396; SHIGARICIN.
DR	SMART: SMO0458; RICIN: 2.
DR	PROSITE: PS50231; RICIN_B_LLECTIN: 2.
KW	Hydrolase; Signal; Toxin.
FT	SIGNAL
FT	1
FT	32
FT	580
FT	CHAIN
SO	SEQUENCE 580 AA; 64421 MW; 940D10F01E7FB558 CRC64;

Query Match	48.9%	Score 1492;	DB 10;	Length 580;
Best Local Similarity	55.3%	Pred. No. 2.6e-113;		
Matches 307;	Conservative 78;	Mismatches 142;	Indels 28;	Gaps 12;

[illegible]

RESULT 3		
ID	PRELIMINARY:	PRT: 580 AA.
OC94BW4		
AC	OC94BW4	
DT	01-DEC-2001 (TREMBLrel. 19, Created)	
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)	
DE	Type 2 ribosome-inactivating protein in cinnamomil II precursor (EC 3.2.2.22) (rRNA N-glycosylase).	
OS	Cinnamomum camphora (Camphor tree).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.	
OX	NCBI_TaxID=13429;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Yang O., Gong Z.Z., Liu W.Y.;	
RT	"Molecular cloning of three type 2 RIP (ribosome-inactivating protein)	
RT	genes encoding cinnamomil proteins and study of their expression	
RT	patterns.";	

RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AY039802; AAK82459.1; -  
 DR InterPro: IPR000772; Rictin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Rictin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICTIN.2.  
 DR PROSITE: PS50231; RICTIN\_B\_LECTIN; 2.  
 DR Hydrolyase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 580  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN I  
 SQ SEQUENCE 580 AA; 64265 MW; 37E4289ECC6C0BF CRC64;  
 Query Match 48.7%; Score 1487; DB 10; Length 580;  
 Best Local Similarity 55.1%; Pred. No. 6,7e-113;  
 Matches 306; Conservative 80; Mismatches 141; Indels 28; Gaps 12;  
 QY 41 YPIINFTGATVOSTNFIRAVRGRLTGADVAHEIPVLPNRVGLPIINQRTIVELSN- 99  
 DB 33 YQVITTTKNAKTSTQFIETALRAQLASGEE-PHGIPVNRDGSITVPDSKREIIVELSNW 91  
 QY 100 HAEISVTLADVTNAYVVGVRAGNSAYFFHPNOEDAEATHTLFDVONRYTFAGGNYD 159  
 DB 92 AADSPVALADVDTNAYVAVYRTGSSQSFRLREDND--PAIENLDPDK-RITFPSSGTY 148  
 QY 160 RLEQAGNLRENIELGNGLEAISAALYYSSTGGTOLPLARSTTICOMISEARFOYT 219  
 DB 149 DLERVAGELREELIGMDPLENAISAL--WTSNLNOORALASLIIVLIQWAVEAVERFEI 206  
 QY 220 EGEHRTIRYNRASAPPSVITLNSWGRSTAIQESNO-GAFASPIOLORNSKFSY 278  
 DB 207 EYRVRESITRAEMFRDPAMLSLENKMSALNAVOOSNOGVSSEVEL--RSISNKPVY 264  
 QY 279 DVSI---LPIITALMYRCAP---PPSQF---SLIRPVVNF-----NADVCMDP 321  
 DB 265 VGSVSDRVISGLAIMLFICRSDRTSSDOFIDHLMIRPLIVDAVATDADNDTCA DP 324  
 QY 322 EPIRYIVGRNGLCVDVDRGFRHNGNAIOLMPCKSNTDANQMTLKRNTIRNSKCLTTY 381  
 DB 325 EPIYRISGRNGLCVDVDRGFRHNGNAIOLMPCKSNTDANQMTLKRNTIRNSKCLTTY 384  
 QY 382 GYSFGVYVMIDYCNATATDATRMQIMDNGTIIINPSSSLVLAATSGNSGTTLYOTNITYAV 441  
 DB 385 GYSAGDYVMIDYCNATATDATRMQIMDNGTIIINPSSSLVLAATSGNSGTTLYOTNITYAV 444  
 QY 442 SOGWLPTNNTOPFTVTIVGLYGLCLQANSQGVWIEDCSSERAKDQWALYADGSIIRPOOR 501  
 DB 445 ROGWLGNNTPEFVTSIVGFNDLCQANGDAMWVECESSKAEQKALYDPGSIIRPHQDR 504  
 QY 502 DNCLTS-DSNIRETVVYIISGSPASSGORWMEKNDGIIILYSGVLVDVADSPSLKQII 560  
 DB 505 DRCLTSDNDSOGSIITISCSGSGESGRWVEFNMNDGTTILNKLGLVADVGNPSLHOII 564  
 QY 561 LLYPLHGDNPQIWLPL 575  
 DB 565 IWPATGKPNQOWMLPL 579  
 RESULT 4  
 Q94BW5 PRELIMINARY; PRT; 581 AA.  
 AC 094BW5;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin I precursor  
 DE (EC 3.2.2.22) (RNA N-glycosidase).  
 OS Cinnaomum camphora (Camphor tree).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnaomum.  
 RX NCBL Taxid=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang O., Gong Z.Z., Liu W.Y.  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnaomolin proteins and study of their expression  
 RT patterns.";  
 RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AY039801; AAK82458.1; -  
 DR InterPro: IPR000772; Rictin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Rictin\_B\_lectin; 5.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICTIN.2.  
 DR PROSITE: PS50231; RICTIN\_B\_LECTIN; 2.  
 DR Hydrolyase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 581  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN I  
 SQ SEQUENCE 581 AA; 64215 MW; 6E8F5FB8FBA3D196 CRC64;  
 Query Match 47.3%; Score 1443.5; DB 10; Length 581;  
 Best Local Similarity 53.8%; Pred. No. 2,4e-109;  
 Matches 299; Conservative 79; Mismatches 149; Indels 29; Gaps 12;  
 QY 41 YPIINFTGATVOSTNFIRAVRGRLTGADVAHEIPVLPNRVGLPIINQRTIVELSN- 99  
 DB 33 YQVITTTKNAKTSTQFIETALRAQLASGEE-PHGIPVNRDGSITVPDSKREIIVELSNW 91  
 QY 100 HAEISVTLADVTNAYVVGVRAGNSAYFFHPNOEDAEATHTLFDVONRYTFAGGNYD 159  
 DB 92 AADSPVALADVDTNAYVAVYRTGSSQSFRLREDND--PAIENLDPDK-RITFPSSGTY 148  
 QY 160 RLEQAGNLRENIELGNGLEAISAALYYSSTGGTOLPLARSTTICOMISEARFOYT 219  
 DB 149 DLERVAGELREELIGMDPLENAISAL--WTSNLNOORALASLIIVLIQWAVEAVERFEI 206  
 QY 220 EGEHRTIRYNRASAPPSVITLNSWGRSTAIQESNO-GAFASPIOLORNSKFSY 278  
 DB 207 EYRVRESITRAEMFRDPAMLSLENKMSALNAVOOSNOGVSSEVEL--RSISNKPVY 264  
 QY 279 DVSI---LPIITALMYRCAP---PPSQF---SLIRPVVNF-----NADVCMDP 321  
 DB 265 VGSVSDRVISGLAIMLFICRSDRTSSDOFIDHLMIRPLIVDAVATDADNDTCA DP 324  
 QY 322 EPIRYIVGRNGLCVDVDRGFRHNGNAIOLMPCKSNTDANQMTLKRNTIRNSKCLTTY 381  
 DB 325 EPIYRISGRNGLCVDVDRGFRHNGNAIOLMPCKSNTDANQMTLKRNTIRNSKCLTTY 384  
 QY 382 GYSFGVYVMIDYCNATATDATRMQIMDNGTIIINPSSSLVLAATSGNSGTTLYOTNITYAV 441  
 DB 385 GYSAGDYVMIDYCNATATDATRMQIMDNGTIIINPSSSLVLAATSGNSGTTLYOTNITYAV 444  
 QY 442 SOGWLPTNNTOPFTVTIVGLYGLCLQANSQGVWIEDCSSERAKDQWALYADGSIIRPOOR 501  
 DB 445 ROGWLGNNTPEFVTSIVGFNDLCQANGDAMWVECESSKAEQKALYDPGSIIRPHQDR 504  
 QY 502 DNCLTS-DSNIRETVVYIISGSPASSGORWMEKNDGIIILYSGVLVDVADSPSLKQII 560  
 DB 505 DRCLTSDNDSOGSIITISCSGSGESGRWVEFNMNDGTTILNKLGLVADVGNPSLHOI 564  
 QY 560 ILYPLHGDNPQIWLPL 575  
 DB 565 IWPATGKPNHMLPL 580

RESULT 5



DT	01-JUN-2002 (TREMBlrel. 21, Created)
DT	01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE	Lectin chain A isoform 1 (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).
OS	Viscum album (European mistletoe).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC	Santalales; Viscaceae; Viscum.
ON	NCBI_TaxID=3972;
RP	SEQUENCE FROM N.A.
RA	TISSUE=Leaf;
RC	Paramasivam M., Misra V., Srinivasan A., Singh T.P.;
RT	"Viscum album (Indian) mRNA for Mistletoe lectin chain A isoform 1 and
RT	chain B."
RL	Submitted (FEF8-2002) to the EMBL/GenBank/DBJ databases.
CC	-1- CATALYTIC ACTIVITY: ENDOLYOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC	SPECIFIC ADENOSINE ON THE 265' RNA.
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR	EMBL: AY081114.9; AA487006.1; "
DR	InterPro: IPR000772; RICH_B_Lectin.
DR	InterPro: IPR001574; RIF.
DR	Pfam: PF00161; RICH_B_Lectin; 6.
DR	Pfam: PF00161; RIF; 1.
DR	SMART: SM00458; RICHIN; 2.
DR	PROSITE: PS50231; RICHIN_B_Lectin; 2.
KW	Hydrolase; Toxin.
FT	NON_TER
FT	CHAIN
FT	CHAIN
SO	SEQUENCE
Query Match	43.1%; Score 1315; DB 10; Length 531;
Best Local Similarity	51.2%; Pred. No. 6.9e-99;
Matches 274; Conservative 82; Mismatches 159; Indels 20; Gaps 11	
QY	48 TAGATVQSYNFRFAVRGRITGADVREHIEYLRNRGGLPIN-QRFITVELSNHAELSV 105
DB	9 THQTGEERFRFTLLINDYVSGS-FSENIPL-RQSTIPVSDQRVLELTNEGDSI 66
QY	106 TLADVTNAYVVGVRAGNSAVFPHDNOEDAETHLEFTVQNRYYTFAFGNDRLEOLA 165
DB	67 TAAIDVTNYYVAYAQDQSYFLR-DVRGAE--THLEFOT-TRSLIPFGSIPDLERA 122
QY	166 GNAREVIELGNGLPELAISALYYTSGTQLPTLARSFTIICOMISDAARFOYIEGEMT 225
DB	123 GH-RDQIPGLDQLQSYVALKE--PGGNTRTORARSLILLOMISEARFNPIIMBARQ 178
QY	226 RIRYNRRSAPDVSITLLENSMGRSLTAQESNQCAFAPSIOLRRNGSPSYDLSILP 285
DB	179 YINSGASFLPDYVMELEFTSMGQSTQVQOSTBEVFENPRLAIPGNEVYTLNARDYA 238
QY	286 IIALMYRCARPSSQ----FSLIRPVYENFNAV-CMDPEIVIRVGRNGLCYVDVRG 340
DB	239 SLAIMEVYGERSSDVRVPLVIRIYADV--DYTGSASEPTVIRVGRNMCVDRBD 295
QY	341 REFHNGNALQWPECKSNTDANQLMLTKRDNTRNSGKLTYYGSPVYVMIYDCATATD 400
DB	296 DEFDGNOQLQWPECKSNTDANQLMLTKRDNTRNSGKLTYYGSPVYVMIYDCATATD 355
QY	401 ATRRQITWDNGLTIINRSSLVATSGNSGTTLTVOGNIYAVSOGWLPNTNNOPTVYTVG 460
DB	356 ATRRQITWDNGLTIINRSSLVATSGNSGTTLTVOGNIYAVSOGWLPNTNNOPTVYTVG 415
QY	461 LVGLICLQNSQGVIEDCSSEKAQWALYADSPTRPQNRNCLTDSNIRREYVTKLS 520
DB	416 FLDLCMESNGSVVWEVCSSQQRALKGDSIPRKQNOQCLTGRDVSIVTINVS 475
QY	521 CGPASGGRMMFKNDGITLNLISGLVLDVRAQSDPSLKLQIILYPLHGDNPQIMLP 575
DB	476 CSAGSGGRMMFKNDGITLNLISGLVLDVRAQSDPSLKLQIILYPLHGDNPQIMLP 530

## RESULT 9

08W243 PRELIMINARY; PRT: 565 AA.  
 AC 08W243;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, last annotation update)  
 DE VCA precursor (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Viscum album subsp. coloratum.  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Santalales; Viscaceae; Viscum.  
 NCBI\_TaxID=159976;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA Park W.-B., Lyu S.;  
 RT Cloning of Viscum album subsp. coloratum (Korean mistletoe).";  
 RL Blochm. Biophys. Res. Commun. 0:0-0(2002).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AF369961; AL40417.1; -;  
 DR InterPro: IPR000772; R1cin\_B\_lectin.  
 DR InterPro: IPR001574; R1P.  
 DR Pfam: PF00652; R1cin\_B\_lectin; 6.  
 DR Pfam: PF00161; R1P; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LLECTIN; 2.  
 KM Hydrolase; Signal; Toxin.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 273 VCA ALPHA CHAIN.  
 FT CHAIN 309 565 VCA BETA CHAIN.  
 SQ SEQUENCE 565 AA; 62401 MW; 991E394DA005F11 CRC64;

Query Match 42.2%; Score 1286; DB 10; Length 565;

Best Local Similarity 49.0%; Pred. No. 1.8e-96;  
 Matches 280; Conservative 85; Mismatches 177; Indels 30; Gaps 14;

15 VAWLCFGSGTSGMFLIEDNNIPKQYPIINFTTACVQSYNFTIRAVAGRLTTGADVR 74  
 12 VSCFLMLGLVGAIVKAEKFRERLRRLVTHQTG--DEYRFTLLDQVSSGS-FS 67  
 75 HEIPLARVAGLPIN--ORFTLVLSNHAELSTYLAADVNAVVGVRAGNSAFYFH--P 130  
 68 NEIPLR-ROSTIPVSADQRFYVELTNOGGSDITRAIDVNLVVAQAQDSQFRLDAP 126  
 131 DNOEDAEATITLFDVONRYTFAFGNDRLEQLAGNRENIENELGNPLEAISALYYYS 190  
 127 DCAE-----RHLEFETG-TRSSLPFTGTYDLERYAGH-RDQIPGLIELIQSVSA- 178  
 191 TGGTOLPTLARSFFIICIMISEARFOYIEGEMFRIRYNRSAPDPSVITLNSMGRLS 250  
 179 GSTR--AQARSLITLQIMISEARFNTFWRARQYINSGESFLPDWYMLELHESWQOS 236  
 251 TAIQESNOGAFASPIQLORNRGSKFSYDVSILPTIALMYRCAPPPSSO----FSLLI 306  
 237 TOVQOSTGCVFNNPRLGISTGNFVTLISNRDVIASLAIMLFVCRDPSSSDVYKPLVI 296  
 307 RPYVNFNA--DY-CMDEPIVIRYGRNGLCVDVDDGFRHNGNAIOLMPCSKSNTDANOLM 363  
 297 RPYLNSGAVDDVDTCTASEPVRIVIRGRGLCVADGKFFYNGNPLOLP-----WDPQOLM 352  
 364 TLKRDNTIRNSNGKCLTYGYSPGYVMYIDCNTAATDTRQIOWDNGTIINPRSLVIAA 423  
 353 TLKRDNTIRNSNGKCLTYGYSPGYVMYIDCNTAATDTRQIOWDNGTIINPRSLVIAA 412  
 424 TSGNSGTTLVQNTINIVASQGMPTNNQOPFVTTIVGLYGLICLQANSQVWIEQSSSEKA 483  
 413 ASSGSGTTLVQVYVFLGQWLAGNDTPAPREVITYGFGNLCMEANGASVSVEVTCGGSKE 472  
 484 EQGMALYADGSTRPQONRNCLESDNIRFETVYKILISGPPASSQGRMKFKNDGTILNLYS 543

Db 473 NOKMALIXGSGTRPQONDOCTTSGGDSVSTVFIVNIVCSAGSSQGRWEFTNEGTTILNAN 532  
 QY 544 GLVLDVNRASDPSLKOITLPLHGDNPQNTMPL 575  
 Db 533 GLVMDVAQSNPSLRIRIITPATGKPNOMLVP 564

## RESULT 10

09M6E9 PRELIMINARY; PRT: 547 AA.  
 AC 09M6E9;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)  
 DE Preproagglutinin (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Agg.  
 OS Abrus precatorius (Indian licorice) (Crib's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eustosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 NCBI\_TaxID=4816;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA MEDLINE=20102702; PubMed=10636890;  
 RA Liu C.-L., Tsai C.C., Lin S.C., Wang L.I., Hsu C.I., Hwang M.J.,  
 RA Lin J.Y.;  
 RT "Primary Structure and Function Analysis of the Abrus precatorius  
 Agglutinin A Chain by Site-directed Mutagenesis: Pro199 of Amphiphilic  
 alpha-Helix H Impairs Protein Synthesis Inhibitory Activity.";  
 RL J. Biol. Chem. 275:1897-1901(2000).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AF190173; AAF28309.1; -;  
 DR HSP: P11140; IABR.  
 DR InterPro: IPR000772; R1cin\_B\_lectin.  
 DR InterPro: IPR001574; R1P.  
 DR Pfam: PF00652; R1cin\_B\_lectin; 6.  
 DR Pfam: PF00161; R1P; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LLECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KM Hydrolase; Toxin.  
 SQ SEQUENCE 547 AA; 61248 MW; 355A325C2354A1BD CRC64;

Query Match 40.8%; Score 1244; DB 10; Length 547;

Best Local Similarity 47.4%; Pred. No. 4.6e-93;  
 Matches 263; Conservative 87; Mismatches 179; Indels 26; Gaps 11;

31 LEDNNIPKQYPIINFTTACVQSYNFTIRAVAGRLTTGADVRHEIPLV--PNRVGLP 87  
 10 LHNAYVYAOQFDPKFTTGSATPASTNORITDLARELTG--LIIGIPLRDPSTVERK 67  
 88 INORFTLVLSNHAELSTYLAADVNAVVGVRAGNSAFYFHNDQEDAEATITLFDVQ 147  
 68 --NQVYVELSVDVSTIOLGIDLTNAVVAVRAGSSSFF--RNAPASATYLTGTGQ 122  
 148 NRYTFAFGNDRLEQLAGNRENIENELGNPLEAISALVYYSTGQTLTLARSFII 207  
 123 -OYSLPFGNVDDEKMAHOSRORISG--LELRGCTIFLRSAGSADDEIARTLIVII 178  
 208 QMISEARFOYIEGEMFRIRYNRSAPDPSVITLNSMGRLSTAIOESNOGAFASPIQL 267  
 179 QMVADEARFRYYSKIVSLISNRARFQDPDSEMLEMTEPLSRAVOHTVODTFPQAVTL 238  
 268 -----QRNGSKFSYDVSILPTIALMYRCAPPPSSQSLIRPVYVFNMDVCDP- 321  
 239 INVQGERVYVSSLSHPSVSAL---ALMLFVCNPLNATQSPDLLIRSYVE--QSKTCSHY 292  
 322 EPIVIRVNGLCVDVDDGFRHNGNAIOLMPCSKSNTDANOLMTLTKRNTIRNSNGKCLTY 381



Db 293 EPTVIRIGRDLGVSDNAYNNGNPILLMKCKDQLEVNOLMTLKSOKTIRSKKCLTTY 352  
 QY 382 GSPGVYMIYDCNTAATDARTMOIWDGTTIINRSSLYLAATSGNSTLTIVONTIYAV 441  
 Db 353 GYABGNMYMIYDCSSAAVEATYWDIWDGTTIINRSGLVLSAESSSMGKTLTVOKNDYRM 412  
 QY 442 SGMPLTNPTOPFVTTIYGLGLQANSGQVMIEDCSSEKAEQOMALYADGSTRPOONR 501  
 Db 413 RQGWRTGNDTSPEVTSINGFEKLCMEAHGNSMMLDVEDYDITKEQOMAYVPDGSIRPVONT 472  
 QY 502 DNCILTSQSNIRETYVKILSCGPASSGQRMFKNDGTTILINYSGLVLDVRAADPSLKOIIL 561  
 Db 473 NNCLTCEHKGATIVMGCSNMAWASQRMVFKSDGTYINLYDVMYVKSQSPSLKQIIL 532  
 QY 562 YPLHGPNOIWLPLF 576  
 Db 533 MPYGNANQMMATLF 547

## RESULT 11

Q8W2E7 PRELIMINARY; PRT; 592 AA.  
 AC Q8W2E7;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Ribosome-inactivating protein IRab (EC 3.2.2.22) (rRNA  
 DE N-glycosidase).  
 GN LECIRAB.  
 OS Iris hollandica (Dutch iris).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;  
 OC Iris.  
 OX NCBI\_TaxID=35876;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Van Damme E.J.M., Peumans W.J.;  
 RT "Iris (Iris hollandica var. Professor Blaauw) plants express both type  
 RT 1 and type 2 ribosome-inactivating proteins in bulb tissue."  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: A6256085; AAL55094.1;  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; R1P.  
 DR Pfam: PF00652; R1cin\_B\_lectin; 6.  
 DR PRINTS: PF00161; R1P; 1.  
 DR SMART: SM00458; SHIGARICIN.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Toxin.  
 SQ SEQUENCE 592 AA; 65677 MW; 6FLP35ECF8A99E3 CRC64;

Query Match 36.3%; Score 1109; DB 10; Length 592;  
 Best Local Similarity 39.6%; Pred. No. 5,7e-82;

Matches 241; Conservative 101; Mismatches 196; Indels 70; Gaps 13;

QY 8 IVIMVAVATWLCFGSTGSMSTLEDDNIFPKOYPIINTTATGATVOSTNFRVAVRGL 67  
 Db 16 IVIMVAVATWLCFGSTGSMSTLEDDNIFPKOYPIINTTATGATVOSTNFRVAVRGL 67  
 QY 68 TGGADVRRHPIVLPNVRGLPT-----NORFIVEISNHAELSVTLADVTNVAVGY 119  
 Db 74 SSGT-----SEYGIPLMRAONPSSSOELLVEIFGMNEPVTILVNVAYLAY 123  
 QY 120 RAGNSAYFFH--PDNOE-----DAEALTHFTVDVQNRVYFAFGCNDRLDEQLAGNREINE 173  
 Db 124 QAGQHYYLLHDPDPNPOLYGSDA-----HRLSFGSYPALQHVAGFEYREND 170  
 QY 174 LONGPLEAIALYVYTGCGTOLPTIARSFIICOMISFAAFQYIEGEMPTRIYRN-- 230

Db 171 LGNELGAILVLHOMSP-PTVERTARSEFVLIQWVSEARERAIE---TRVRNRIQ 225  
 QY 231 ----RSAPDPVYITLNSWGRSLSTAIOESNOGAFASPIQORNGSKSVYDVALIRI 286  
 Db 226 VGDYRSRPRPAGMLDETNNGTLSERVQESNEGVFNRLTLQTTNETTHIYNAQTAROV 285  
 QY 287 --TALVYRCAPPSSQPSL-----LIRVPNPNFNAVCMDEPVIYRVG 329  
 Db 286 CGLALLLFCARQSIALPLPHDSVPLPLLDLVNVRMSMLDVEDDTPCPEPPTIRISG 345  
 QY 330 RNLGVADVRRGRPHNCAIQIOWPCNKSTNDANDLTLKBNRTIRSNCKCLTTYGSPGVY 389  
 Db 346 RDGYCMQVADGILHGNPNVTLSSCKONVNOFWFKSGTIOISNCKCLTAVYAGAVY 405  
 QY 390 MIYDCNTAATDARTMOIWDGTTIINRSSLYLAATSGNSTLTIVONTIYAVSOGMLPNP 449  
 Db 406 MIYDCSSAVVAGATLMY-NGSLINRPSGLAISAESEGGTTLTMQVHLNAKQKQVSN 464  
 QY 450 NTQPFVTTIYGLGLQANSGQ-VWIEDCSSEKAEQOMALYADGSTRPOONRDNCLTSD 508  
 Db 465 NTRPFLPTIINGLGVORNDQEDVGLATCDNNNSQKMYLYGDSIRPLTPNVCVYSQ 524  
 QY 509 SNIRETVYKILSCGPASSGQRMFKNDGTTILINYSGLVLDVRAADPSLKOIILYPLHGP 568  
 Db 525 THQGSQIILLSNCFGASQRMVFTSQGTYINLYHSGYVADVKQSDPSLQIITWSTGNP 584  
 QY 569 NOIWLPLF 576  
 Db 585 NQMMFTTF 592

## RESULT 12

Q8W2E8 PRELIMINARY; PRT; 573 AA.  
 AC Q8W2E8;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Ribosome-inactivating protein IRAR (EC 3.2.2.22) (rRNA N-glycosidase)  
 DE (fragment).  
 GN LECIRAR.  
 OS Iris hollandica (Dutch iris).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;  
 OC Iris.  
 OX NCBI\_TaxID=35876;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Van Damme E.J.M., Peumans W.J.;  
 RT "Iris (Iris hollandica var. Professor Blaauw) plants express both type  
 RT 1 and type 2 ribosome-inactivating proteins in bulb tissue."  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: A6256084; AAL55093.1;  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; R1P.  
 DR Pfam: PF00652; R1cin\_B\_lectin; 6.  
 DR PRINTS: PF00161; R1P; 1.  
 DR SMART: SM00458; SHIGARICIN.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Toxin.  
 FT NON\_TER  
 SQ SEQUENCE 573 AA; 63759 MW; 1414A3B9AECDAF5C CRC64;

Query Match 35.4%; Score 1081; DB 10; Length 573;  
 Best Local Similarity 40.0%; Pred. No. 1.1e-79;

Matches 230; Conservative 101; Mismatches 176; Indels 68; Gaps 13;

QY 41 YPIINFTTATGATVOSTNFRVAVRGLTGGADVRRHPIVLPNVRGLPT-----NORF 92







•  
•  
•  
•

PT N-terminal methionine free proteins prodn. - by using host

```

XX  Disclosure; Fig. 4; 20pp; English.
PS
XX  Ricin A may be produced in a form which lacks an N-terminal Met
CC  using Met-aminopeptidase from E.coli.
XX
SQ  Sequence 332 AA;

Query Match      100.0%; Score 948; DB 8; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.6e-92;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  1 MKPGGNTIVIMYAAVATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTTAGATVOSTYNTFI 60
    1 MKPGGNTIVIMYAAVATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTTAGATVOSTYNTFI 60
DB  61 RAVRGRLTTGADVREHIEIVLPNVRGVLPIINORFIVELSNHAEISVTLADVTNAYVGYR 120
    61 RAVRGRLTTGADVREHIEIVLPNVRGVLPIINORFIVELSNHAEISVTLADVTNAYVGYR 120
OY  121 AGNSAYFFHPDNOEDAEATITHEFTDVONRYTFAGGNYDRLEOLAGNLRNIEELGNGPL 179
    121 AGNSAYFFHPDNOEDAEATITHEFTDVONRYTFAGGNYDRLEOLAGNLRNIEELGNGPL 179
DB  121 AGNSAYFFHPDNOEDAEATITHEFTDVONRYTFAGGNYDRLEOLAGNLRNIEELGNGPL 179

RESULT 2
AAP70838
ID  AAP70838 standard; protein; 332 AA.
XX
AC  AAP70838;
XX
DT  25-MAR-2003 (updated)
DT  18-FEB-1991 (first entry)
XX
DE  Sequence of Ricinus communis castor beans ricin toxin (Rt or ricin)
XX  A protein encoded by PRA123.
XX
KW  Lectin; toxin protein; cytotoxic; cytostatic; castor bean;
KW  plant toxin.
XX
OS  Ricinus communis.
XX
FH  Key
FT  Region 1..32 Location/Qualifiers
FT  Region 33..302 /note="Leader"
FT  Region /note="A-chain"
FT  Region 315..332 /note="B-chain"
FT  Region /note="B-chain"
XX
PN  EP237676-A.
XX
PD  23-SEP-1987.
XX
PF  13-NOV-1986; 86EP-0308877.
XX
PR  07-MAR-1986; 86US-0837583.
XX
PA  (CETU ) CETUS CORP.
PA  (CHIR ) CHIRON CORP.
XX
PI  Plateak M;
XX
DR  WPI; 1987-265177/38.
DR  N-PSDB; AAN70519.
XX
PT  New non-glycosylated ricin precursor and toxin etc. - are prepd.
PT  by recombinant DNA procedures with specific isolation steps for
PT  pure and soluble prods.
XX
PS  Disclosure; Fig 1; 112pp; English.
XX
CC  The full-length sequences encoding ricin A (AAN70520), ricin D

```

```

CC  (AAN70525), putative ricin E (AAN70526) and RCA (AAN70524) in precursor
CC  form were obt'd. using messenger RNA to obtain a cDNA library, and
CC  then probing the library to retrieve the desired cDNA inserts. The
CC  library was probed using the 35-mer given in AAN70514. Figure 4 (see
CC  AAN70520, AAN70521, AAN70522) shows the nucleotide sequences of three
CC  plasmids contg. cDNA inserts obt'd. by probing a cDNA library for
CC  sequences encoding ricin B using the probe in AAN70517. The cDNA
CC  inserts can be placed into expression vectors. Site-directed
CC  mutagenesis may be used to place an AUG start codon and a HindIII
CC  site at the beginning of the mature protein (see AAN70518). The
CC  coding sequences of the inserts can be ligated into expression
CC  vectors contg. the Phoa promoter-operator and leader sequence
CC  (AAN70523) and suitable retroregulators.
XX  (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ  Sequence 332 AA;

Query Match      100.0%; Score 948; DB 8; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.6e-92;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  1 MKPGGNTIVIMYAAVATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTTAGATVOSTYNTFI 60
    1 MKPGGNTIVIMYAAVATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTTAGATVOSTYNTFI 60
DB  61 RAVRGRLTTGADVREHIEIVLPNVRGVLPIINORFIVELSNHAEISVTLADVTNAYVGYR 120
    61 RAVRGRLTTGADVREHIEIVLPNVRGVLPIINORFIVELSNHAEISVTLADVTNAYVGYR 120
OY  121 AGNSAYFFHPDNOEDAEATITHEFTDVONRYTFAGGNYDRLEOLAGNLRNIEELGNGPL 179
    121 AGNSAYFFHPDNOEDAEATITHEFTDVONRYTFAGGNYDRLEOLAGNLRNIEELGNGPL 179
DB  121 AGNSAYFFHPDNOEDAEATITHEFTDVONRYTFAGGNYDRLEOLAGNLRNIEELGNGPL 179

RESULT 3
AAP95639
ID  AAP95639 standard; protein; 332 AA.
XX
AC  AAP95639;
XX
DT  25-MAR-2003 (updated)
DT  31-OCT-2002 (updated)
DT  13-AUG-1990 (first entry)
XX
DE  Ricin A encoded by insert from plasmid PRA123.
XX
KW  Plasmid PRA123; ricin-A; ricin-B; cytotoxicity.
XX
OS  Ricinus communis.
XX
OS  Synthetic.
XX
FH  Key
FT  Peptide 1..35 Location/Qualifiers
FT  Peptide /label= Leader sequence
FT  Peptide 36..302 /label=A-chain
FT  Peptide 303..314 /label=linker
FT  Peptide 315..332 /label=B-chain
FT  Peptide /label=B-chain
XX
PN  EP335476-A.
XX
PD  04-OCT-1989.
XX
PE  19-JAN-1989; 89EP-0201162.
XX
PR  08-FEB-1984; 84US-0578115.
PR  08-FEB-1984; 84US-0578121.
PR  09-FEB-1984; 84US-0578122.
PR  07-SEP-1984; 84US-0648759.
PR  20-SEP-1984; 84US-0653515.
XX

```

PA (CETU ) CETUS CORPORATION.  
 XX Gelfand D, Lawyer FC, Horn G, Greenfield L, Nliecekl D, Kaplan D;  
 PI Platak MJ;  
 XX WPI: 1989-286959/40.  
 DR N-PSDB: AAN91281.  
 XX  
 PT Recombinant vectors expressing ricin chains or diphtheria toxin -used for  
 PT Prod. of new immunotoxin conjugates with monoclonal antibodies, having  
 PT high cell specificity and good extracellular stability.  
 XX  
 PS Disclosure; Fig 14; 54pp; English.  
 CC  
 CC Plasmid RA123 (ATCC No. 39799) carries the entire coding sequence for  
 CC ricin A, as well as codons for 12 AAs joining the A to the B chain.  
 CC Following modification for ease of manipulation the plasmid was used to  
 CC construct expression vectors which express the conjugates in  
 CC host cells.  
 CC (Updated on 31-OCT-2002 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 332 AA:  
 Query Match 100.0%; Score 948; DB 10; Length 332;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-92;  
 Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKPGGNTIYIMWYAVATWLCFSGTSGMSFTLEDNNIFPKOYPIINFTTAGATVQSYTNFI 60  
 DB 1 MKPGGNTIYIMWYAVATWLCFSGTSGMSFTLEDNNIFPKOYPIINFTTAGATVQSYTNFI 60  
 QY 61 RAVRGRLTGADVREHIEPVLPRVGLPINORFIVELSNHAEISVTALDVTNAYVGYR 120  
 DB 61 RAVRGRLTGADVREHIEPVLPRVGLPINORFIVELSNHAEISVTALDVTNAYVGYR 120  
 QY 121 AGNSAYFFHPDNOEDAETHLFTDVQNRRTFAFGGNYDRLEQLAGNIRENIELGNGPL 179  
 DB 121 AGNSAYFFHPDNOEDAETHLFTDVQNRRTFAFGGNYDRLEQLAGNIRENIELGNGPL 179  
 RESULT 4  
 AAP90079 standard; protein; 562 AA.  
 XX  
 AC AAP90079;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Ricin D.  
 XX  
 KW Ricin D; Ricinus communis; castor beans; Zaubariensis variety;  
 KW modified; lectin binding removed; reduced cell binding  
 XX  
 OS Ricinus communis (castor beans).  
 XX  
 PN NO8904839-A.  
 XX  
 PD 01-JUN-1989.  
 XX  
 PF 23-NOV-1988; 88WO-US04238.  
 XX  
 PR 24-NOV-1987; 87US-0124735.  
 XX  
 PA (GEMY ) GENETICS INST INC.  
 XX  
 PI Brown EL, Jones S;  
 XX  
 DR WPI: 1989-178366/24.  
 DR N-PSDB: AAN90068.  
 DR

XX  
 PT Modified ricin molecules and toxin conjugates  
 PT - in which the lectin binding function of the B chain  
 PT is removed or diminished to reduce cell binding.  
 XX  
 PS Disclosure; fig 1; 51pp; English.  
 CC  
 CC Ricin D (see corresp. AAN90068) found on the EcoRI-HindIII fragment  
 CC of DNA from Ricinus communis, Zaubariensis variety. Patent  
 CC discloses many modifications of ricin in which the lectin binding  
 CC function of the B chain is diminished or removed, and conjugation  
 CC to toxins to eliminate cell binding.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 562 AA:  
 Query Match 100.0%; Score 948; DB 10; Length 562;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-92;  
 Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKPGGNTIYIMWYAVATWLCFSGTSGMSFTLEDNNIFPKOYPIINFTTAGATVQSYTNFI 60  
 DB 1 MKPGGNTIYIMWYAVATWLCFSGTSGMSFTLEDNNIFPKOYPIINFTTAGATVQSYTNFI 60  
 QY 61 RAVRGRLTGADVREHIEPVLPRVGLPINORFIVELSNHAEISVTALDVTNAYVGYR 120  
 DB 61 RAVRGRLTGADVREHIEPVLPRVGLPINORFIVELSNHAEISVTALDVTNAYVGYR 120  
 QY 121 AGNSAYFFHPDNOEDAETHLFTDVQNRRTFAFGGNYDRLEQLAGNIRENIELGNGPL 179  
 DB 121 AGNSAYFFHPDNOEDAETHLFTDVQNRRTFAFGGNYDRLEQLAGNIRENIELGNGPL 179  
 RESULT 5  
 AAP70326 standard; Protein: 576 AA.  
 XX  
 ID AAP70326 standard; Protein: 576 AA.  
 XX  
 AC AAP70326;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 21-MAY-1991 (first entry)  
 XX  
 DE Sequence of Ricinus communis (castor bean) Ricin toxin  
 DE (RT or ricin) E precursor encoded by PR38.  
 XX  
 KW Lectin; toxin protein; cytotoxic; cytostatic; castor bean;  
 KW plant toxin.  
 XX  
 OS Ricinus communis.  
 XX  
 FH Key  
 FT Region  
 FT Region  
 FT Region  
 FT Region  
 FT Region  
 XX  
 PN EP237676-A.  
 XX  
 PD 23-SEP-1987.  
 XX  
 PF 13-NOV-1986; 86EP-0308877.  
 XX  
 PR 07-MAR-1986; 86US-0837583.  
 XX  
 PA (CETU ) CETUS CORP.  
 PA (CHTR ) CHIRON CORP.  
 XX  
 PI Platak M;  
 XX  
 DR WPI: 1987-265177/38.  
 DR

DR N-PSDB; AAN70526.  
 XX  
 PT New non-glycosylated ricin precursor and toxin etc. - are prepd.  
 PT by recombinant DNA procedures with specific isolation steps for  
 PT purer and soluble prods.  
 XX  
 PS Disclosure; Fig 14(1-2); 112pp; English.  
 XX  
 CC The full length sequences encoding ricin A (AAN70520), ricin D  
 CC (AAN70525) putative ricin E (AAN70526) and RCA (AAN70524) in precursor  
 CC form were obtained, using the messenger RNA to obtain a cDNA library, and  
 CC then probing the library to retrieve the desired cDNA inserts. The  
 CC library was probed using the 35-mer given in AAN70514. Figure 4 (see  
 CC AAN70520, AAN70521, AAN70522). Shows the nucleotide sequences of three  
 CC plasmids containing cDNA inserts obtained by probing a cDNA library  
 CC for sequences encoding ricin B using the probe in AAN70517. The cDNA  
 CC inserts can be placed into expression vectors. Site-directed  
 CC mutagenesis may be used to place an ATG start codon and a HindIII  
 CC site at the beginning of the mature protein. (see AAN70518). The  
 CC coding sequences of the inserts can be ligated into expression  
 CC vectors containing the Phoa promoter-operator and leader sequence  
 CC (AAN70523) and suitable retroregulators.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 XX  
 SQ Sequence 576 AA:  
 Query Match 100.0%; Score 948; DB 8; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-92;  
 Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKPGGNTIVIMWYAAVATWLCFGSTSGWSTFLENNIFPKOYPIINFTTAGATVQSYTNFI 60  
 DB 1 MKPGGNTIVIMWYAAVATWLCFGSTSGWSTFLENNIFPKOYPIINFTTAGATVQSYTNFI 60  
 QY 61 RAVRGRLTGGADVRRHEIPVLPNRVGLPINORFILVELSNHAEISVTLALDVTNAYVGYR 120  
 DB 61 RAVRGRLTGGADVRRHEIPVLPNRVGLPINORFILVELSNHAEISVTLALDVTNAYVGYR 120  
 QY 121 AGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAFGNDRLEQLAGNLENIELGNPL 179  
 DB 121 AGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAFGNDRLEQLAGNLENIELGNPL 179  
 RESULT 6  
 AAN25787  
 ID AAN25787 standard; Protein; 576 AA.  
 XX  
 AC AAN25787;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 27-MAR-1998 (first entry)  
 XX  
 DE Castorbean ricin.  
 XX  
 KW Ricin; cytotoxin; hybrid protein; cell delivery;  
 KW cell binding ligand; translocation domain; diphtheria toxin B';  
 KW interleukin-2; T-cell lymphoma; organ rejection; therapy.  
 XX  
 OS Ricinus communis.  
 XX  
 FH key Location/Qualifiers  
 FT Peptide 1..35  
 FT /label= sig\_peptide  
 FT Protein 36..302  
 FT /label= A-domain  
 FT Peptide 303..314  
 FT /label= linker  
 FT Domain 315..576  
 FT /label= B-domain  
 XX  
 XX US5668255-A.  
 XX 16-SEP-1997.  
 PD

XX  
 PF 04-AUG-1993; 93US-0102387.  
 XX  
 PR 27-JUN-1991; 91US-0722484.  
 PR 07-JUN-1984; 84US-0618199.  
 PR 25-APR-1985; 85US-0726808.  
 PR 07-JUN-1985; 85US-0742554.  
 PR 22-DEC-1989; 89US-0456095.  
 PR 14-JUN-1990; 90US-0538276.  
 PR 04-AUG-1993; 93US-0102387.  
 XX  
 PA (SERA-) SERAGEN INC.  
 XX  
 PI Murphy JR.  
 XX  
 DR WPI: 1997-470103/43.  
 DR N-PSDB; AAT91638.  
 XX  
 PT New hybrid molecules for delivery of agents to cells - comprise a  
 PT binding domain of a cell binding ligand and a portion of a  
 PT translocation domain of a protein  
 XX  
 PT Example 4; Fig 11A-B; 30pp; English.  
 XX  
 PS This polypeptide comprises the castorbean cytotoxin, ricin.  
 XX DNA (see AAT91638) encoding the enzymatic A domain and a portion  
 CC of the A-to-B linker peptide of ricin was used to construct a  
 CC ricin-diphtheria toxin B'-interleukin-2 gene that was expressed in  
 CC E. coli. The hybrid protein can be isolated and used to treat  
 CC conditions involving over-production of cells bearing IL2 receptors,  
 CC such as certain T-cell lymphomas and organ transplant rejection  
 CC crises. The hybrid inactivates ribosomes in cells bearing IL2  
 CC receptors, resulting in cessation of protein synthesis and death of  
 CC target cells. Claimed hybrid proteins comprise a translocation  
 CC domain and a cell binding domain from e.g. a hormone, growth factor  
 CC or polypeptide toxin. The hybrid molecules can be used for the  
 CC delivery of agents (e.g. therapeutic genes, toxins, detectable  
 CC labels) into cells. The use of a translocation mechanism ensures  
 CC that the hybrid will be effective in relatively low doses, since a  
 CC high proportion of the substance of interest will be taken into the  
 CC targeted cells. The hybrid molecules can be manufactured as a  
 CC single hybrid recombinant protein, permitting reproducibility,  
 CC consistency, and the precise control of composition.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC  
 XX  
 SQ Sequence 576 AA:  
 Query Match 100.0%; Score 948; DB 18; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-92;  
 Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKPGGNTIVIMWYAAVATWLCFGSTSGWSTFLENNIFPKOYPIINFTTAGATVQSYTNFI 60  
 DB 1 MKPGGNTIVIMWYAAVATWLCFGSTSGWSTFLENNIFPKOYPIINFTTAGATVQSYTNFI 60  
 QY 61 RAVRGRLTGGADVRRHEIPVLPNRVGLPINORFILVELSNHAEISVTLALDVTNAYVGYR 120  
 DB 61 RAVRGRLTGGADVRRHEIPVLPNRVGLPINORFILVELSNHAEISVTLALDVTNAYVGYR 120  
 QY 121 AGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAFGNDRLEQLAGNLENIELGNPL 179  
 DB 121 AGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAFGNDRLEQLAGNLENIELGNPL 179  
 RESULT 7  
 AAY55892  
 ID AAY55892 standard; Protein; 576 AA.  
 XX  
 AC AAY55892;  
 XX  
 DT 15-FEB-2000 (first entry)  
 DT Castor bean ricin toxin.  
 XX  
 DE

XX Recombinant; hybrid; binding domain; ligand; animal cell; diphtheria;  
 KW translocation domain; botulinum; neurotoxin; ricin; cholera; tetanus;  
 KW shiga-like toxin; pertussis; translocation; cytoplasmic membrane; HIV;  
 KW cytosol; therapy; genetic deficiency disease; enzyme; co-factor; poison;  
 KW adipocyte; cancer; virus; infection; antibody.

XX Ricinus communis.

XX US9565406-A.

XX 12-OCT-1999.

XX 07-JUN-1995; 95US-0488246.

XX 04-AUG-1993; 93US-0102387.

XX 07-JUN-1984; 84US-0618199.

XX 27-JUN-1991; 91US-0722484.

XX 25-APR-1985; 85US-0726808.

XX 07-JUN-1985; 85US-0742554.

XX 22-DEC-1989; 89US-0456095.

XX 14-JUN-1990; 90US-0538276.

XX (SERA-) SERAGEN INC.

XX Murphy JR;

XX WPI: 1999-632431/54.

XX N-PDB: AA230663.

XX Example 4; Fig 11; 31pp; English.

XX The invention relates to a recombinant DNA molecule encoding a hybrid  
 CC protein comprising three parts: (a) the first part comprises a portion  
 CC of the binding domain of a cell-binding polypeptide ligand allowing the  
 CC hybrid protein to bind to an animal cell; (b) the second part comprises  
 CC a portion of a translocation domain of a naturally occurring protein  
 CC selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera  
 CC toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus  
 CC membrane into the cytosol of the cell; and (c) the third part comprises  
 CC a polypeptide entity to be introduced into the cell, which is non-native  
 CC to the naturally occurring protein of (b). This sequence represents the  
 CC Castor bean ricin toxin sequence for use in generating the hybrid of the  
 CC invention. The hybrid molecule enables the direction of appropriate  
 CC therapy to affected cells, allowing them to function properly and  
 CC alleviate or cure the disease. The hybrid is especially used in treating  
 CC genetic deficiency diseases, by delivering to affected cells an enzyme  
 CC supplying the missing function, to supplementing cellular levels of a  
 CC particular enzyme or a scarce precursor or cofactor, to directing toxins  
 CC or other poisons to destroy particular cells (such as adipocytes, cancer  
 CC cell, or virus infected cells), to counteracting viral infections such as  
 CC HIV, by introducing appropriate antibodies to viral proteins. It is also  
 CC involved in the process of getting non-therapeutic substances such as  
 CC detectable labels into cells.

XX Sequence 576 AA;

XX Query Match

XX Best Local Similarity 100.0%; Score 948; DB 20; Length 576;

XX Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 MKRGAGTIVIMAVATWLCFGSTSGMFLLEDNNIFPKOYPIINFTAGATVOSTYNTI 60

XX 1 MKRGAGTIVIMAVATWLCFGSTSGMFLLEDNNIFPKOYPIINFTAGATVOSTYNTI 60

XX 61 RAVRGRITGADVREHIEPVPNVRGPIPNORFTVLSNHAELSTVLAIDYNAVYVGR 120

XX 61 RAVRGRITGADVREHIEPVPNVRGPIPNORFTVLSNHAELSTVLAIDYNAVYVGR 120

OY 121 AGNSAVFFHPDNOEDAEATHTLFTDVONRYTFAFGGNDRLEQLAGNIRENTELGNGPL 179  
 DB 121 AGNSAVFFHPDNOEDAEATHTLFTDVONRYTFAFGGNDRLEQLAGNIRENTELGNGPL 179

RESULT 8

AA78592

AA78592 standard; Protein; 576 AA.

AA78592;

05-MAY-2000 (first entry)

Ricinus communis ricin protein sequence.

Ricin; toxin; hybrid protein; translocation domain; cell destruction;  
 KW cell binding domain; genetic deficiency disease; cell targeting; cancer;  
 KW adipocyte; enzyme delivery; anti-viral; HIV.

OS Ricinus communis.

US6022950-A.

08-FEB-2000.

07-JUN-1995; 95US-0479510.

07-JUN-1984; 84US-0618199.

27-JUN-1991; 91US-0722484.

25-APR-1985; 85US-0726808.

07-JUN-1985; 85US-0742554.

22-DEC-1989; 89US-0456095.

14-JUN-1990; 90US-0538276.

04-AUG-1993; 93US-0102387.

(SERA-) SERAGEN INC.

Murphy JR;

WPI: 2000-160390/14.

N-PDB: AA290019.

Example 4; Fig 11; 32pp; English.

XX This sequence represents the Ricinus communis ricin protein sequence. The  
 CC toxin can be included in the hybrid protein of the invention and used to  
 CC destroy or modify the cell that the hybrid protein is targeted to. The  
 CC hybrid protein comprises a first part which is a portion of the binding  
 CC domain of a cell-binding ligand, effective to cause the hybrid molecule  
 CC to bind to a cell of an animal. The second part comprises a portion of a  
 CC translocation domain of a naturally occurring protein (e.g. the  
 CC translocation domain of diphtheria toxin) the second part translocates  
 CC the third part across the cytoplasmic membrane and into the cytosol of  
 CC the cell. The third part comprises a chemical entity to be introduced  
 CC into the cell, where each of the first and third part is non-native with  
 CC respect to naturally occurring protein, and the covalent bond attaching  
 CC the second and third part is cleavable. The toxin represented by the  
 CC present sequence can form part of the third portion of the hybrid  
 CC protein. The cell binding domain binds to a specific cell and the  
 CC translocation domain transfers the hybrid molecule across the cell  
 CC membrane into the cytosol. The third part of the protein, linked to the  
 CC translocation domain through a cleavable bond, can then carry out its  
 CC function. The hybrid molecules are useful for treating genetic deficiency  
 CC diseases by delivering to affected cells an enzyme supplying the missing  
 CC function, to supplement cellular levels of a particular enzyme or a  
 CC scarce precursor or cofactor, to direct toxins or other poisons to  
 CC destroy particular cells (such as adipocytes, cancer cells, or  
 CC virus-infected cells), and to counteract viral infections such as HIV by  
 CC introducing into appropriate cells antibodies to viral proteins.

```

XX Sequence      576 AA:
SQ
Query Match      100.0%; Score 948; DB 21; Length 576;
Best Local Similarity 100.0%; Pred. No. 3.2e-92;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKPGNTIVIMYAVATWLCFGSTSGSFLEDDNNIFPKQYPIINFAGATVQSYTNFI 60
    |||
DB 1 MKPGNTIVIMYAVATWLCFGSTSGSFLEDDNNIFPKQYPIINFAGATVQSYTNFI 60
    |||

OY 61 RAVGRLLTGADVRHEIPLPNRVGLPINORFIVELSNHAEISVTALADVTNAYVGYR 120
    |||
DB 61 RAVGRLLTGADVRHEIPLPNRVGLPINORFIVELSNHAEISVTALADVTNAYVGYR 120
    |||

OY 121 AGNSAVFFHPDNOEDAEATHTFTDVONRRTFAFGNDRLEOLAGNLRNIEELGNGPL 179
    |||
DB 121 AGNSAVFFHPDNOEDAEATHTFTDVONRRTFAFGNDRLEOLAGNLRNIEELGNGPL 179
    |||

RESULT 9
AAG78301 standard; Protein: 576 AA.
XX AC AAG78301;
XX DT 15-NOV-2001 (first entry)
XX DE Castor bean preproricin protein (SEQ ID 2).
XX KW Castor bean plant; preproricin; ricin; A chain; B chain;
    human immunodeficiency virus infection; HIV; toxin; antiviral agent;
    retroviral infection; anti-HIV; virucide activity; viral protease.
XX OS Ricinus communis.
XX FH Key Location/Qualifiers
    FT Peptide 1..35
        /label= Signal peptide
    FT Protein 36..302
        /label= Ricin_A_chain
        /note= "N-glycosidase"
    FT Peptide 303..314
        /label= Linker_peptide
    FT Protein 315..576
        /note= "Cleaved during activation of ricin"
    FT /label= Ricin_B_chain
    FT /note= "Galactose/N-acetylglactosamine-binding lectin"
XX OS
XX PN WO200160393-A1.
XX PD 23-AUG-2001.
XX PF 15-FEB-2001; 2001WO-US05282.
XX PR 16-FEB-2000; 2000US-0182759.
XX PA (BECH-) BECHTEL BWXT IDAHO LLC.
XX PI Keener WK, Ward TE;
XX DR WPI; 2001-581908/65.
XX DR N-PSDB; AA164138.
XX PT Novel composition comprising toxin e.g., ricin based antiviral compound
    useful for treating viral infections such as human immunodeficiency
    virus infection.
XX PS Disclosure; Page 50-54; 66pp; English.
XX CC The sequence relates to preproricin protein encoded by the DNA sequence
    given in AA164138. The invention relates to a novel toxin (e.g., ricin)
    based antiviral agent which is toxic to virus-infected cells, but

```

```

CC non-toxic to uninfected cells. The invention has anti-HIV and virucide
CC activities. Its mechanism of action is through inactivation of cellular
CC ribosomes and enhancement of binding of the antiviral agent to galactose
CC residues on cell surfaces, and its cellular internalisation. The
CC invention is useful for treating human immunodeficiency virus infection
CC and other viral infections, especially retroviral infections. The
CC antiviral agent is activated in viral particles or early-stage infected
CC cells, killing the cells upon infection and effectively preventing the
    integration of the viral genome into the host genome thereby preventing
    the latency/rebound problem. The agent enters all HIV susceptible cells,
    and not just cells known to act as host cells for the virus. The
    cell is infected with the virus, where the viral protease activates it.
XX
SQ Sequence      576 AA:
Query Match      100.0%; Score 948; DB 22; Length 576;
Best Local Similarity 100.0%; Pred. No. 3.2e-92;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKPGNTIVIMYAVATWLCFGSTSGSFLEDDNNIFPKQYPIINFAGATVQSYTNFI 60
    |||
DB 1 MKPGNTIVIMYAVATWLCFGSTSGSFLEDDNNIFPKQYPIINFAGATVQSYTNFI 60
    |||

OY 61 RAVGRLLTGADVRHEIPLPNRVGLPINORFIVELSNHAEISVTALADVTNAYVGYR 120
    |||
DB 61 RAVGRLLTGADVRHEIPLPNRVGLPINORFIVELSNHAEISVTALADVTNAYVGYR 120
    |||

OY 121 AGNSAVFFHPDNOEDAEATHTFTDVONRRTFAFGNDRLEOLAGNLRNIEELGNGPL 179
    |||
DB 121 AGNSAVFFHPDNOEDAEATHTFTDVONRRTFAFGNDRLEOLAGNLRNIEELGNGPL 179
    |||

RESULT 10
AAG78302 standard; Protein: 576 AA.
XX ID AAG78302;
XX AC AAG78302;
XX DT 15-NOV-2001 (first entry)
XX DE Castor bean preproricin protein (SEQ ID 3).
XX KW Castor bean plant; preproricin; ricin; A chain; B chain;
    human immunodeficiency virus infection; HIV; toxin; antiviral agent;
    retroviral infection; anti-HIV; virucide activity; viral protease.
XX OS Ricinus communis.
XX FH Key Location/Qualifiers
    FT Peptide 1..35
        /label= Signal peptide
    FT Protein 36..302
        /label= Ricin_A_chain
        /note= "N-glycosidase"
    FT Peptide 303..314
        /label= Linker_peptide
    FT Protein 315..576
        /note= "Cleaved during activation of ricin"
    FT /label= Ricin_B_chain
    FT /note= "Galactose/N-acetylglactosamine-binding lectin"
XX OS
XX PN WO200160393-A1.
XX PD 23-AUG-2001.
XX PF 15-FEB-2001; 2001WO-US05282.
XX PR 16-FEB-2000; 2000US-0182759.
XX PA (BECH-) BECHTEL BWXT IDAHO LLC.
XX PI Keener WK, Ward TE;

```



XX WPI: 2001-581908/65.  
 DR N-PSDB: AAI64139.  
 XX Novel composition comprising toxin e.g., ricin based antiviral compound  
 PT useful for treating viral infections such as human immunodeficiency  
 PT virus infection.  
 XX  
 PS Disclosure; Page 54-57; 66pp; English.  
 XX  
 CC The sequence relates to preproricin protein encoded by the DNA sequence  
 CC given in AAI64139. The invention relates to a novel toxin (e.g., ricin)  
 CC based antiviral agent which is toxic to virus-infected cells, but  
 CC non-toxic to uninfected cells. The invention has anti-HIV and virucide  
 CC activities. Its mechanism of action is through inactivation of cellular  
 CC ribosomes and enhancement of binding of the antiviral agent to galactose  
 CC residues on cell surfaces, and its cellular internalisation. The  
 CC invention is useful for treating human immunodeficiency virus infection  
 CC and other viral infections, especially retroviral infections. The  
 CC antiviral agent is activated in viral particles or early-stage infected  
 CC cells, killing the cells upon infection and effectively preventing the  
 CC integration of the viral genome into the host genome thereby preventing  
 CC the latency/rebound problem. The agent enters all HIV susceptible cells,  
 CC and not just cells known to act as host cells for the virus. The  
 CC antiviral agent remains inert in a cell until degraded in it, unless the  
 CC cell is infected with the virus, where the viral protease activates it.  
 CC  
 XX  
 SQ Sequence 576 AA:  
 Query Match 100.0%; Score 948; DB 22; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-92;  
 Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKPGNTIVIMYAVATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTTAGATVOSTYTNFI 60  
 Db 1 MKPGNTIVIMYAVATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTTAGATVOSTYTNFI 60  
 QY 61 RAVGRLLTGADVRHEIPVLPNRYGLPINORFLIVELSNHAELSVTLADVTNAYVVGXR 120  
 Db 61 RAVGRLLTGADVRHEIPVLPNRYGLPINORFLIVELSNHAELSVTLADVTNAYVVGXR 120  
 QY 121 AGNSAFEFHPDNOEDAIHTLFTDVONRYTFAGGNYDRLEOLAGNLRENIELGNGPL 179  
 Db 121 AGNSAFEFHPDNOEDAIHTLFTDVONRYTFAGGNYDRLEOLAGNLRENIELGNGPL 179  
 RESULT 11  
 AAR70325  
 ID AAR70325 standard; protein; 574 AA.  
 XX  
 AC AAR70325;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 18-FEB-1991 (first entry)  
 XX  
 DE Sequence of Ricinus communis castor beans ricin toxin (RT or ricin)  
 DE D precursor encoded by pRT17.  
 XX  
 KM Lectin: toxin protein; cytotoxic; cytostatic; castor bean;  
 KM plant toxin.  
 XX  
 OS Ricinus communis.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..33  
 FT /note="Leader"  
 FT Region 34..300  
 FT /note="A-chain"  
 FT Region 313..574  
 FT /note="B-chain"  
 XX  
 PN EP237676-A.  
 XX

PD 23-SEP-1987.  
 XX  
 PF 13-NOV-1986; 86EP-0308877.  
 XX  
 PR 07-MAR-1986; 86US-0837583.  
 XX  
 PA (CETU ) CETUS CORP.  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Piatek M;  
 XX  
 XX WPI: 1987-265177/38.  
 DR N-PSDB: AAN70525.  
 XX  
 PT New non-glycosylated ricin precursor and toxin etc. - are prepd.  
 PT by recombinant DNA procedures with specific isolation steps for  
 PT pure and soluble prods.  
 XX  
 PS Disclosure; Fig 13(1-2); 112pp; English.  
 XX  
 CC The full-length sequences encoding ricin A (AAN70520), ricin D  
 CC (AAN70525), putative ricin E (AAN70526) and RCA (AAN70524) in precursor  
 CC form were obtd. using messenger RNA to obtain a cDNA library, and  
 CC then probing the library to retrieve the desired cDNA inserts. The  
 CC library was probed using the 35-mer given in AAN70514. Figure 4 (see  
 CC AAN70520, AAN70521, AAN70522) shows the nucleotide sequences of three  
 CC plasmids contg. cDNA inserts obtd. by probing a cDNA library for  
 CC sequences encoding ricin B using the probe in AAN70517. The cDNA  
 CC inserts can be placed into expression vectors. Site-directed  
 CC mutagenesis may be used to place an ATG start codon and a HindIII  
 CC site at the beginning of the mature protein (see AAN70518). The  
 CC coding sequences of the inserts can be ligated into expression  
 CC vectors contg. the Phoa promoter-operator and leader sequence  
 CC (AAN70533) and suitable retroregulators.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 574 AA:  
 Query Match 98.9%; Score 938; DB 8; Length 574;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-91;  
 Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PGCGNTIVIMYAVATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTTAGATVOSTYTNFIRA 62  
 Db 1 PGCGNTIVIMYAVATWLCFGSTSGMSFTLEDNNIFPKOYPIINFTTAGATVOSTYTNFIRA 60  
 QY 63 VRGRLLTGADVRHEIPVLPNRYGLPINORFLIVELSNHAELSVTLADVTNAYVVGXRAG 122  
 Db 61 VRGRLLTGADVRHEIPVLPNRYGLPINORFLIVELSNHAELSVTLADVTNAYVVGXRAG 120  
 QY 123 NSAYFEFHPDNOEDAIHTLFTDVONRYTFAGGNYDRLEOLAGNLRENIELGNGPL 179  
 Db 121 NSAYFEFHPDNOEDAIHTLFTDVONRYTFAGGNYDRLEOLAGNLRENIELGNGPL 177  
 RESULT 12  
 AAR06554  
 ID AAR06554 standard; protein; 332 AA.  
 XX  
 AC AAR06554;  
 XX  
 DT 10-MAR-2003 (updated)  
 DT 04-JAN-1991 (first entry)  
 XX  
 DE Ricin A gene product of plasmid insert pRA123.  
 XX  
 KM Alkaline phosphatase: phoA;  
 KM Ricinus communis.  
 XX  
 OS Ricinus communis.  
 OS  
 PN US948729-A.  
 XX  
 PD 14-AUG-1990.  
 XX

```

XX 17-MAR-1988; 88US-0171132.
PF 17-MAR-1988; 88US-0171132.
XX 25-MAR-1985; 83US-0715933.
PR 25-MAR-1985; 83US-0715933.
XX (CETU ) CETUS CORP.
PA (CETU ) CETUS CORP.
XX Platak M, Laird WJ, Lane JA;
PI WPI: 1990-267900/35.
XX N-PSDB; AAQ05786.
DR WPI: 1990-267900/35.
XX Soluble, biologically active protein prodn. in prokaryotic hosts
PT - using terminated leader DNA sequence encoding bacterial
PS alkaline phosphatase
XX Disclosure; Fig 8; 19pp; English.
XX Pho A leader sequence can be used to manufacture genes encoding
CC soluble, active, heterologous proteins for a bacterial expression
CC system. Heterologous ricin A is soluble in the sonicate of cells
CC and may be readily purified without detergent.
CC (updated on 10-MAR-2003 to add missing OS field.)
XX SQ Sequence 332 AA;
SQ Query Match 98.8%; Score 937; DB 11; Length 332;
Best Local Similarity 98.9%; Pred. No. 2.3e-91;
Matches 177; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKPGNTIYIMYAVATWLCFSGSTGMSFTLEDNNIPFKOYPIINFAGATVOSTYTFI 60
DB 1 MKPGNTIYIMYAVATWLCFSGSTGMSFTLEDNNIPFKOYPIINFAGATVOSTYTFI 60
QY 61 RAVRGRLTGADVRHREIPVLPNRVGLPINOREFIVELSNHAEISVTALADVTNAYVGYR 120
DB 61 RAVRGRLTGADVRHREIPVLPNRVGLPINOREFIVELSNHAEISVTALADVTNAYVGYR 120
QY 121 AGNSAYFFHPDQEDAEATHTLFTDVONRYTFAPFGNTDRLEOLAGNIRENIEIGNGPL 179
DB 121 AGNSAYFFHPDQEDAEATHTLFTDVONRYTFAPFGNTDRLEOLAGNIRENIEIGNGPL 179

RESULT 13
AAP94793
ID AAP94793 standard; protein; 574 AA.
XX AAP94793;
AC AAP94793;
XX 06-JUL-1990 (first entry)
DT 06-JUL-1990 (first entry)
XX DNA sequence of ricin toxin D.
DE DNA sequence of ricin toxin D.
XX Ricin; toxin; baculovirus; mutants; ss.
KW Ricin; toxin; baculovirus; mutants; ss.
XX Baculovirus.
OS Baculovirus.
XX Key 1.35 Location/Qualifiers
FH Peptide /label=Leader peptide.
FT Domain 36..500
FT Domain /label=A-chain
FT Domain /label=B-chain
FT Domain 301..574
FT Domain /label=B-chain
XX W08901037-A.
PN W08901037-A.
XX 09-FEB-1989.
PD 09-FEB-1989.
XX 20-JUL-1988; 88WO-US02442.
PF 20-JUL-1988; 88WO-US02442.
XX 24-JUL-1987; 87US-0077126.
PR 24-JUL-1987; 87US-0077126.

```

```

PR 08-FEB-1988; 88US-0153778.
XX (CETU ) CETUS CORP.
PA (CETU ) CETUS CORP.
XX Houston LL, Lane JA, Platak M, Clark R;
PI Houston LL, Lane JA, Platak M, Clark R;
XX WPI: 1989-061173/08.
DR WPI: 1989-061173/08.
XX N-PSDB; AAN91039.
DR N-PSDB; AAN91039.
XX Recombinant baculovirus transfer vectors -
PT Recombinant baculovirus transfer vectors -
PT used for prodn. of ricin toxin in a baculovirus insect cell
PT expression system.
XX Disclosure; Page -; 85pp; English.
XX Recombinant baculovirus vector is capable of introducing ricin gene into
CC virus genome, allowing creation of a insect cell/baculovirus ricin
CC expression system.
XX SQ Sequence 574 AA;
SQ Query Match 98.0%; Score 929; DB 10; Length 574;
Best Local Similarity 99.4%; Pred. No. 3.4e-90;
Matches 176; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 PGNTIYIMYAVATWLCFSGSTGMSFTLEDNNIPFKOYPIINFAGATVOSTYTFI 62
DB 1 PGNTIYIMYAVATWLCFSGSTGMSFTLEDNNIPFKOYPIINFAGATVOSTYTFI 60
QY 63 VRGRLTGADVRHREIPVLPNRVGLPINOREFIVELSNHAEISVTALADVTNAYVGYRAG 122
DB 61 VRGRLTGADVRHREIPVLPNRVGLPINOREFIVELSNHAEISVTALADVTNAYVGYRAG 120
QY 123 NSAYFFHPDQEDAEATHTLFTDVONRYTFAPFGNTDRLEOLAGNIRENIEIGNGPL 179
DB 121 NSAYFFHPDQEDAEATHTLFTDVONRYTFAPFGNTDRLEOLAGNIRENIEIGNGPL 177

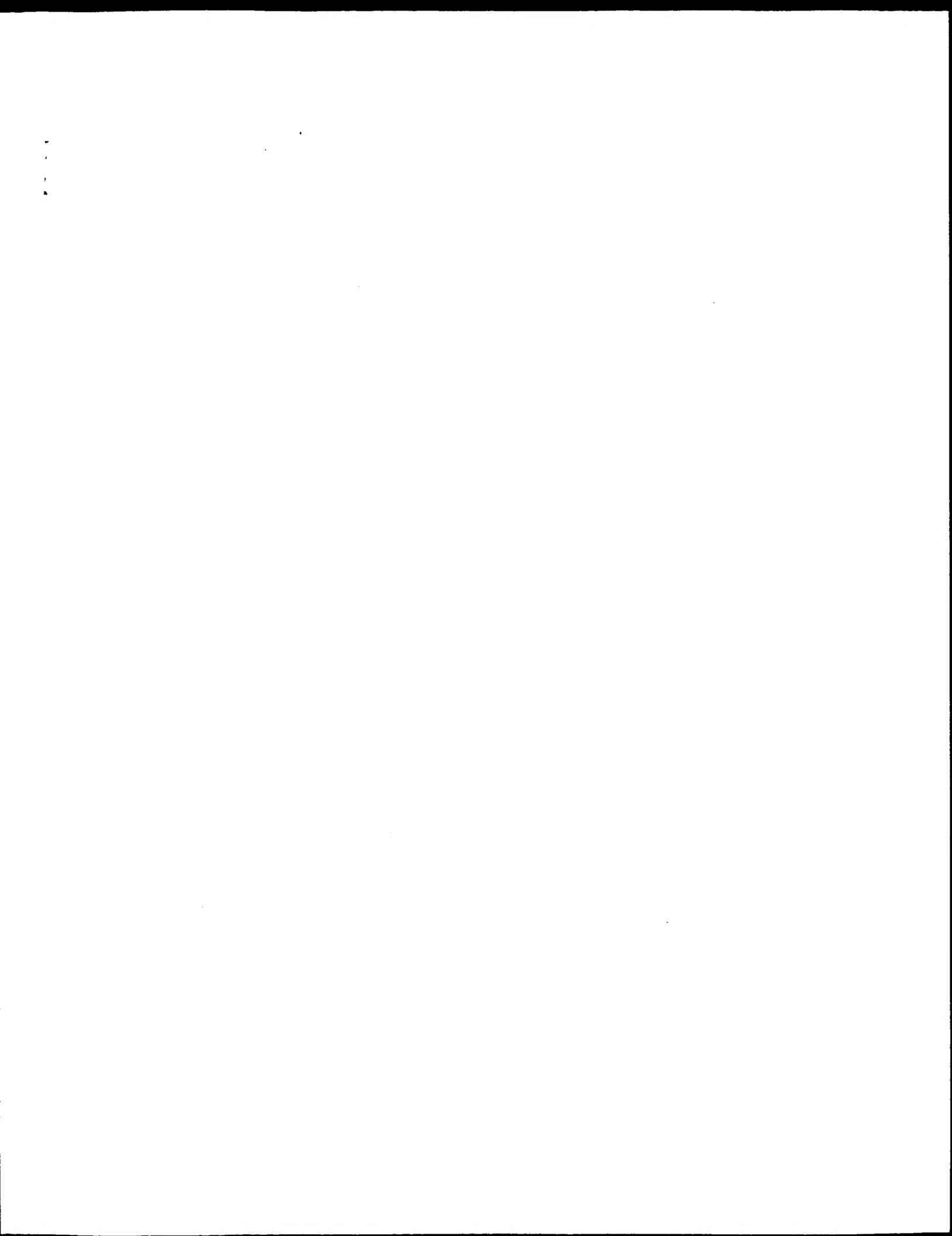
RESULT 14
AAM25136
ID AAM25136 standard; protein; 290 AA.
XX AAM25136;
AC AAM25136;
XX 25-MAR-2003 (updated)
DT 02-DEC-1997 (first entry)
XX Ricin A-chain ribosome inhibitory protein inactive precursor.
DE Ricin A-chain ribosome inhibitory protein inactive precursor.
XX Malze; proRIP; ribosome inactivating protein; alpha; beta subunit;
KW internal linker; barley translation inhibitor; trichosanthin;
KW Ricin A-chain; Abrin-A-chain; Saporin; SLT-1; Luffin A; MAP;
KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;
KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;
KW post-translational modification; cancer; neoplasia; HIV; AIDS;
KW human immunodeficiency virus; acquired immune deficiency syndrome.
XX Synthetic.
OS Synthetic.
XX US5646026-A.
PN US5646026-A.
XX 08-JUL-1997.
PD 08-JUL-1997.
XX 07-JUN-1995; 95US-0485286.
PF 07-JUN-1995; 95US-0485286.
XX 09-DEC-1992; 92US-0987927.
PR 11-JUN-1990; 90US-0535636.
PR 26-JAN-1995; 95US-0378761.
PR 07-JUN-1995; 95US-0485286.
XX (DOWC ) DOWELANCO.
PA (DOWC ) DOWELANCO.
XX Hey TD, Morgan AER, Walsh TA;
PI Hey TD, Morgan AER, Walsh TA;

```

XX WP1: 1997-362934/33.  
 XX DNA encoding pro-ribosome inactivating proteins - inactive  
 PT precursors of ribosome inactivating proteins; can be expressed in  
 PT eukaryotic cells without causing cell death  
 XX  
 PS Claim 4: Column 91-94; 186pp; English.  
 XX  
 CC AAM25136 represents a Ricin A-chain ribosome inhibitory protein (RIP)  
 CC which was engineered to contain a selectively removable internal peptide  
 CC linker sequence separating the alpha and beta units of the RIP. When  
 CC separated the two units regain activity and are capable of inactivating  
 CC eukaryotic ribosomes and hence preventing protein production. Many  
 CC different RIPs may be produced with an internal linker including  
 CC maize RIP, Trichosanthin, Ricin A-chain, Abrin-A A-chain and  
 CC Saporin. The RIPs can be used in the construction of therapeutic  
 CC toxins targeted to specific cells such as tumour cells via the  
 CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.  
 CC A further use is in HIV therapy (see US4869903). There is interest  
 CC in expressing RIP recombinantly in host eukaryotic cells, because of  
 CC the capacity to provide correct post-translational processing. However,  
 CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting  
 CC in cell death. Since the inactive RIP proteins are not cytotoxic to  
 CC eukaryotic cells, they can be recombinantly expressed in such cells and  
 CC then converted to active RIP proteins.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 290 AA;  
 QY  
 Query Match 93.4%; Score 885; DB 18; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 6,8e-86;  
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 MYAVATWLCFGSGTSGMSFTLEDNNIPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGA 71  
 1 MYAVATWLCFGSGTSGMSFTLEDNNIPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGA 60  
 QY 72 DVRRHEIPVLPNRRVGLPIINQRFILVELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPD 131  
 61 DVRRHEIPVLPNRRVGLPIINQRFILVELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPD 120  
 Db 132 NOEDAEATHTLFTDVQNRRTFAFGNVDRLDQLAGNIRENIELGNGPL 179  
 121 NOEDAEATHTLFTDVQNRRTFAFGNVDRLDQLAGNIRENIELGNGPL 168  
 QY  
 Db 132 NOEDAEATHTLFTDVQNRRTFAFGNVDRLDQLAGNIRENIELGNGPL 179  
 121 NOEDAEATHTLFTDVQNRRTFAFGNVDRLDQLAGNIRENIELGNGPL 168  
 RESULT 15  
 AAM21699  
 ID AAM21699 standard; Protein; 290 AA.  
 XX  
 AC AAM21699;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 26-SEP-1997 (first entry)  
 XX  
 DE Ricin A-chain RIP.  
 XX  
 KW pro-Ribosome inactivating Protein; proRIP; peptide linker; cancer;  
 KW inactivation; eukaryotic ribosome; alpha fragment; beta fragment;  
 KW inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;  
 KW rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.  
 XX  
 OS Ricinus communis.  
 XX  
 FH Key Location/Qualifiers  
 FT 152..162  
 FT Region /note="Position of possible insertion of internal  
 FT peptide linker sequence"  
 XX  
 PN US5635384-A.  
 XX  
 PD 03-JUN-1997.

XX  
 PF 26-JAN-1995; 95US-0378761.  
 XX  
 PR 09-DEC-1992; 92US-0987927.  
 PR 11-JUN-1990; 90US-0535636.  
 PR 26-JAN-1995; 95US-0378761.  
 XX  
 PA (DOMC ) DOWELANCO.  
 XX  
 PI Hey TD, Morgan AER, Walsh TA;  
 XX  
 DR WP1: 1997-309831/28.  
 XX  
 PT Inactive precursor of maize ribosome-inactivating protein - also  
 PT chimeric ribosome-inactivating protein precursors containing  
 PT internal linker sequences  
 XX  
 PS Claim 2: Column 91-94; 121pp; English.  
 XX  
 CC The sequences given in AAM21698-710 represent Ribosome inactivating  
 CC proteins (RIP's), which may be used in the construction of the  
 CC proRIP of the invention. The proRIP has a selectively removable,  
 CC internal peptide linker. The precursor sequence is incapable of  
 CC inactivating eukaryotic ribosomes, but can be converted by removal  
 CC of the linker into a protein having alpha and beta fragments and being  
 CC capable of inactivating eukaryotic ribosomes. RIPs are potent  
 CC inhibitors of eukaryotic protein synthesis. They possess a highly  
 CC specific N-glycosidase activity which cleaves the glycosidic bond of  
 CC adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit  
 CC cellular proliferation of cells, e.g. cancer cells and HIV-infected T  
 CC cells. The inactive proRIP proteins make it possible to provide protein  
 CC synthesis inhibitors with uses in practical and improved ways not before  
 CC possible. The RIP can be used to make cytotoxic conjugates.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 290 AA;  
 QY  
 Query Match 93.4%; Score 885; DB 18; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 6,8e-86;  
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 MYAVATWLCFGSGTSGMSFTLEDNNIPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGA 71  
 1 MYAVATWLCFGSGTSGMSFTLEDNNIPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGA 60  
 QY 72 DVRRHEIPVLPNRRVGLPIINQRFILVELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPD 131  
 61 DVRRHEIPVLPNRRVGLPIINQRFILVELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPD 120  
 Db 132 NOEDAEATHTLFTDVQNRRTFAFGNVDRLDQLAGNIRENIELGNGPL 179  
 121 NOEDAEATHTLFTDVQNRRTFAFGNVDRLDQLAGNIRENIELGNGPL 168  
 QY  
 Db 132 NOEDAEATHTLFTDVQNRRTFAFGNVDRLDQLAGNIRENIELGNGPL 179  
 121 NOEDAEATHTLFTDVQNRRTFAFGNVDRLDQLAGNIRENIELGNGPL 168

Search completed: September 16, 2003, 11:45:16  
 Job time : 30.8072 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:42:03 ; Search time 10.1972 Seconds  
(without alignments)  
742.718 Million cell updates/sec

Title: US-10-083-336A-2

Perfect score: 948  
Sequence: 1 MKPGNTIYIMYAVATWLC.....RLEQLAGNIRENIEIENGPL 179

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/iaa/5a\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5b\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6a\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6b\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	885	93.4	290	1	US-08-378-761A-27
2	885	93.4	290	1	US-08-485-286-27
3	885	93.4	290	6	5248606-4
4	748	78.9	267	1	US-07-901-707-1
5	748	78.9	267	1	US-07-988-430-1
6	748	78.9	267	1	US-08-425-336-1
7	748	78.9	267	1	US-08-488-113B-1
8	748	78.9	267	1	US-08-477-484B-1
9	748	78.9	267	2	US-08-446-360-1
10	748	78.9	267	3	US-08-839-765-1
11	748	78.9	267	3	US-09-136-389-1
12	748	78.9	267	4	US-09-610-838-1
13	748	78.9	267	5	PCT-US92-09487-1
14	748	78.9	268	2	US-08-356-786-8
15	748	78.9	534	2	US-08-356-786-10
16	738	77.8	267	1	US-08-218-303-16
17	738	77.8	267	2	US-08-318-793D-61
18	738	77.8	267	4	US-09-538-873-1
19	676.5	71.4	540	1	US-08-378-761A-77
20	676.5	71.4	540	1	US-08-485-286-77
21	265.5	28.0	247	1	US-08-488-113B-6
22	265.5	28.0	247	1	US-08-477-484B-6
23	265.5	28.0	247	2	US-08-646-360-6
24	265.5	28.0	247	2	US-08-839-765-6
25	265.5	28.0	247	3	US-09-136-389-6
26	265.5	28.0	247	4	US-09-610-838-6
27	265.5	28.0	267	1	US-08-378-761A-74

28	265.5	28.0	267	1	US-08-485-286-74	Sequence 74, Appl
29	265.5	28.0	289	1	US-07-923-692C-4	Sequence 4, Appl
30	265.5	28.0	289	1	US-08-184-237-4	Sequence 4, Appl
31	265.5	28.0	289	2	US-08-482-920-4	Sequence 4, Appl
32	265.5	28.0	289	3	US-08-484-341-4	Sequence 4, Appl
33	265.5	28.0	289	3	US-08-483-502-4	Sequence 4, Appl
34	265.5	28.0	289	4	US-08-726-651A-4	Sequence 4, Appl
35	261.5	27.6	282	1	US-08-324-301-15	Sequence 15, Appl
36	245	25.8	250	1	US-08-378-761A-71	Sequence 71, Appl
37	245	25.8	250	1	US-08-485-286-71	Sequence 71, Appl
38	245	25.8	251	4	US-09-538-873-3	Sequence 3, Appl
39	240.5	25.4	290	1	US-08-245-754A-2	Sequence 2, Appl
40	240.5	25.4	290	2	US-08-597-731-2	Sequence 2, Appl
41	236.5	24.9	248	3	US-08-902-486-7	Sequence 7, Appl
42	236.5	24.9	496	3	US-08-902-486-15	Sequence 15, Appl
43	234	24.7	255	1	US-07-901-707-6	Sequence 6, Appl
44	234	24.7	255	1	US-07-988-430-6	Sequence 6, Appl
45	234	24.7	255	1	US-08-425-336-6	Sequence 6, Appl

#### ALIGNMENTS

RESULT 1  
US-08-378-761A-27  
Sequence 27, Application US/08378761A  
Patent No. 5635384

#### GENERAL INFORMATION:

APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
TITLE OF INVENTION: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268

#### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/378,761A  
FILING DATE: 26-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T.  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 290 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-378-761A-27

Query Match

Best Local Similarity 93.4%; Score 885; DB 1; Length 290;

Matches 168; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

DB 12 MYAVATWLCFSGTSGMSFTLEDNNIFPKOYIIINFTTAGATVOCYSTYFIIRVGRRLTGA 71  
1 MYAVATWLCFSGTSGMSFTLEDNNIFPKOYIIINFTTAGATVOCYSTYFIIRVGRRLTGA 60



TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-901-707-1

Query Match  
Best Local Similarity 100.0%; Pred. No. 1.7e-78;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTAGATVOSTYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINORFTLY 95  
DB 1 IFPKQYPIINFTAGATVOSTYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINORFTLY 60  
QY 96 ELSNHAELSVTLALDVTNNAVYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVONRTTFAPFG 155  
DB 61 ELSNHAELSVTLALDVTNNAVYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVONRTTFAPFG 120  
QY 156 GNYDRLEQLAGNLRNIEELGNPL 179  
DB 121 GNYDRLEQLAGNLRNIEELGNPL 144

RESULT 5  
US-07-988-430-1  
Sequence 1, Application US/07988430  
Patent No. 5416202  
GENERAL INFORMATION:  
APPLICANT: Bernhardt, Susan L.  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Lane, Julie A.  
APPLICANT: Lei, Shau-Ping  
TITLE OF INVENTION: Materials Comprising and Methods of  
TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/988,430  
FILING DATE: 19921209  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5416202and, Greta E.  
REGISTRATION NUMBER: 35302  
REFERENCE/DOCKET NUMBER: 31133  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-988-430-1

Query Match  
Best Local Similarity 100.0%; Pred. No. 1.7e-78;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTAGATVOSTYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINORFTLY 95  
DB 1 IFPKQYPIINFTAGATVOSTYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINORFTLY 60  
QY 96 ELSNHAELSVTLALDVTNNAVYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVONRTTFAPFG 155  
DB 61 ELSNHAELSVTLALDVTNNAVYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVONRTTFAPFG 120  
QY 156 GNYDRLEQLAGNLRNIEELGNPL 179  
DB 121 GNYDRLEQLAGNLRNIEELGNPL 144

RESULT 6  
US-08-425-336-1  
Sequence 1, Application US/08425336  
Patent No. 5621083  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studulka, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/425,336  
FILING DATE: 18-APR-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Thomas C.  
REGISTRATION NUMBER: P-36,989  
REFERENCE/DOCKET NUMBER: 31394  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-425-336-1

Query Match 78.9%; Score 748; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-78;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINORFTLV 95  
DB 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINORFTLV 60

QY 96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHTLFTDVONRYTFAG 155  
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHTLFTDVONRYTFAG 120

QY 156 GNYDRLEOLAGNLRENIEELGNGPL 179  
DB 121 GNYDRLEOLAGNLRENIEELGNGPL 144

RESULT 7  
US-08-488-113B-1  
; Sequence 1, Application US/08488113B  
; Patent No. 5744580  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488, 113B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-488-113B-1  
Query Match 78.9%; Score 748; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-78;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINORFTLV 95  
DB 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINORFTLV 60

QY 96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHTLFTDVONRYTFAG 155  
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHTLFTDVONRYTFAG 120

QY 156 GNYDRLEOLAGNLRENIEELGNGPL 179  
DB 121 GNYDRLEOLAGNLRENIEELGNGPL 144

RESULT 8  
US-08-477-484B-1  
; Sequence 1, Application US/08477484B  
; Patent No. 5736699  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,484B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids



TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-477-484B-1

Query Match  
Best Local Similarity 100.0%; Score 748; DB 1; Length 267;  
Pred. No. 1.7e-78;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVREIPLPVRVGLPINORFTLY 95  
DB 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVREIPLPVRVGLPINORFTLY 60  
QY 96 ELSNHAELSVTLALDVNTNAYVGYRAGNSAYFFHPDNOEDAIAITHLFTDVQNRRTFAFG 155  
DB 61 ELSNHAELSVTLALDVNTNAYVGYRAGNSAYFFHPDNOEDAIAITHLFTDVQNRRTFAFG 120  
QY 156 GNYDRLSQLAGNLRNIEIENGPL 179  
DB 121 GNYDRLSQLAGNLRNIEIENGPL 144

## RESULT 9

US-08-646-360-1  
Sequence 1, Application US/08646360

PATENT NO. 5837491  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studulka, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/546,360  
FILING DATE: 13-MAY-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-9155  
TELEFAX: 312/707-9155  
TELE: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-360-1

Query Match  
Best Local Similarity 100.0%; Score 748; DB 2; Length 267;  
Pred. No. 1.7e-78;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVREIPLPVRVGLPINORFTLY 95  
DB 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVREIPLPVRVGLPINORFTLY 60  
QY 96 ELSNHAELSVTLALDVNTNAYVGYRAGNSAYFFHPDNOEDAIAITHLFTDVQNRRTFAFG 155  
DB 61 ELSNHAELSVTLALDVNTNAYVGYRAGNSAYFFHPDNOEDAIAITHLFTDVQNRRTFAFG 120  
QY 156 GNYDRLSQLAGNLRNIEIENGPL 179  
DB 121 GNYDRLSQLAGNLRNIEIENGPL 144

## RESULT 10

US-08-839-765-1  
Sequence 1, Application US/08839765

PATENT NO. 6146631  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studulka, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,765  
FILING DATE: 15-APR-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US09/200-70.03.C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-9155  
TELEFAX: 312/707-9155

TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-839-765-1

Query Match 78.9%; Score 748; DB 3; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-78;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFFKQYPIINFTAGATVOSTNFTIRAVGRLLTGADVRHEIPVLPNRVGLPIINORFLV 95  
DB 1 IFFKQYPIINFTAGATVOSTNFTIRAVGRLLTGADVRHEIPVLPNRVGLPIINORFLV 60  
QY 96 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVONRTTFAFG 155  
DB 61 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVONRTTFAFG 120  
QY 156 GNYDRLEQLAGNLRENIEELNGPL 179  
DB 121 GNYDRLEQLAGNLRENIEELNGPL 144

## RESULT 11

US-09-136-389-1  
; Sequence 1, Application US/09136389  
; Patent No. 6146850

GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/136,389  
; FILING DATE:

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/646,360  
; FILING DATE: 13-MAY-1996  
; APPLICATION NUMBER: PCT/US94/05348  
; FILING DATE: 12-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-136-389-1

Query Match 78.9%; Score 748; DB 3; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-78;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFFKQYPIINFTAGATVOSTNFTIRAVGRLLTGADVRHEIPVLPNRVGLPIINORFLV 95  
DB 1 IFFKQYPIINFTAGATVOSTNFTIRAVGRLLTGADVRHEIPVLPNRVGLPIINORFLV 60  
QY 96 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVONRTTFAFG 155  
DB 61 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVONRTTFAFG 120  
QY 156 GNYDRLEQLAGNLRENIEELNGPL 179  
DB 121 GNYDRLEQLAGNLRENIEELNGPL 144

## RESULT 12

US-09-610-838-1  
; Sequence 1, Application US/09610838  
; Patent No. 6376217

GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/610,838  
; FILING DATE: 06-JUL-2000

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/136,389  
; FILING DATE: 18-AUG-1998  
; APPLICATION NUMBER: 08/646,360  
; FILING DATE: 13-MAY-1996  
; APPLICATION NUMBER: PCT/US94/05348  
; FILING DATE: 12-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-9155
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-610-838-1

```

```

Query Match      78.9%; Score 748; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.7e-78;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 36 IFPKQYPIINFTAGATVOSTYNTFIRAVGRLLTGADVHREIPVLPNRYGLPIINORFIV 95
DB 1 IFPKQYPIINFTAGATVOSTYNTFIRAVGRLLTGADVHREIPVLPNRYGLPIINORFIV 60
QY 96 ELSNHAELSVTLALDVNTNAYVVGYSAGNSAYFFHPDNOEDAEATHTLFTDVONRYTPAFG 155
DB 61 ELSNHAELSVTLALDVNTNAYVVGYSAGNSAYFFHPDNOEDAEATHTLFTDVONRYTPAFG 120
QY 156 GNYDRLEQLAGNLRNENIELGNPL 179
DB 121 GNYDRLEQLAGNLRNENIELGNPL 144

```

```

RESULT 13
PCT-US92-09487-1
; Sequence 1, Application PC/TUS9209487
; GENERAL INFORMATION:
; APPLICANT: Bernhard, Susan L.
; APPLICANT: Bette, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Lane, Julie A.
; APPLICANT: Lei, Shau-ping
; TITLE OF INVENTION: Materials Comprising and Methods of
; TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09487
; FILING DATE: 19921104
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:

```

```

NAME: Noland, Greta E.
REGISTRATION NUMBER: 35302
REFERENCE/DOCKET NUMBER: 31133
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-09487-1

```

```

Query Match      78.9%; Score 748; DB 5; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.7e-78;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 36 IFPKQYPIINFTAGATVOSTYNTFIRAVGRLLTGADVHREIPVLPNRYGLPIINORFIV 95
DB 1 IFPKQYPIINFTAGATVOSTYNTFIRAVGRLLTGADVHREIPVLPNRYGLPIINORFIV 60
QY 96 ELSNHAELSVTLALDVNTNAYVVGYSAGNSAYFFHPDNOEDAEATHTLFTDVONRYTPAFG 155
DB 61 ELSNHAELSVTLALDVNTNAYVVGYSAGNSAYFFHPDNOEDAEATHTLFTDVONRYTPAFG 120
QY 156 GNYDRLEQLAGNLRNENIELGNPL 179
DB 121 GNYDRLEQLAGNLRNENIELGNPL 144

```

```

RESULT 14
US-08-356-786-8
; Sequence 8, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids

```

TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-356-786-8

Query March 78.9%; Score 748; DB 2; Length 268;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-78;  
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRIHEIPVLPNVRGLPIINQRTILV 95  
 DB 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRIHEIPVLPNVRGLPIINQRTILV 61  
 QY 96 ELSNHAELSVTLADVTNAYVVGGRAGNSAFFEHPDNOEDAETHLFTDVQNRRTFAFG 155  
 DB 62 ELSNHAELSVTLADVTNAYVVGGRAGNSAFFEHPDNOEDAETHLFTDVQNRRTFAFG 121  
 QY 156 GNYDRLEOLAGNLRENTEELGNGPL 179  
 DB 122 GNYDRLEOLAGNLRENTEELGNGPL 145

## RESULT 15

US-08-356-786-10  
 : Sequence 10, Application US/08356786  
 : Patent No. 5877305

## GENERAL INFORMATION:

APPLICANT: Huston, James S.  
 APPLICANT: Oppermann, Hermann  
 APPLICANT: Houston, L. L.  
 APPLICANT: Ring, David B.  
 TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
 STREET: Exchange Place, 53 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/356,786  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/831,967  
 FILING DATE: 06-FEB-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pitcher, Edmund R.  
 REGISTRATION NUMBER: 27,829  
 REFERENCE/DOCKET NUMBER: CRP-053  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 248-7000  
 TELEFAX: (617) 248-7100  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 534 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-356-786-10

## Query Match

Best Local Similarity 78.9%; Score 748; DB 2; Length 534;  
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRIHEIPVLPNVRGLPIINQRTILV 95

DB 4 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRIHEIPVLPNVRGLPIINQRTILV 63  
 QY 96 ELSNHAELSVTLADVTNAYVVGGRAGNSAFFEHPDNOEDAETHLFTDVQNRRTFAFG 155  
 DB 64 ELSNHAELSVTLADVTNAYVVGGRAGNSAFFEHPDNOEDAETHLFTDVQNRRTFAFG 123  
 QY 156 GNYDRLEOLAGNLRENTEELGNGPL 179  
 DB 124 GNYDRLEOLAGNLRENTEELGNGPL 147

Search completed: September 16, 2003, 11:51:46  
 Job time: 11.1972 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 16, 2003, 11:49:23 ; Search time 18.5118 Seconds

(without alignments)  
1439.723 Million cell updates/sec

Title: US-10-083-336a-2

Perfect score: 948  
Sequence: 1 MKPGGNTIVIMYAVATWLC.....RLEQLAGLRENIEIGNGPL 179

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 556269 seqs, 148893369 residues

Total number of hits satisfying chosen parameters: 556269

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PC1\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PC1\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	748	78.9	267	12	US-10-127-890-1
2	738	77.8	267	12	US-10-282-935-1
3	265.5	28.0	247	10	US-09-792-793A-39
4	265.5	28.0	247	12	US-10-127-890-6
5	265.5	28.0	289	12	US-10-280-679B-4
6	245	25.8	251	10	US-10-282-935-3
7	236.5	24.9	247	10	US-09-792-793A-34
8	210.5	22.2	263	12	US-10-127-890-4
9	206.5	21.8	252	9	US-09-347-064-2
10	206.5	21.6	252	9	US-09-347-064-8
11	184	19.4	248	12	US-10-127-890-7
12	177.5	18.7	251	12	US-10-127-890-5
13	177.5	18.7	251	12	US-10-127-890-110
14	176.5	18.6	251	12	US-10-127-890-111
15	176.5	18.6	251	12	US-10-127-890-101

16	175.5	18.5	251	12	US-10-127-890-99	Sequence 99, App1
17	175.5	18.5	251	12	US-10-127-890-107	Sequence 107, App
18	174.5	18.4	251	12	US-10-127-890-100	Sequence 100, App
19	174.5	18.4	251	12	US-10-127-890-106	Sequence 106, App
20	173.5	18.3	251	9	US-09-765-527-247	Sequence 247, App
21	173.5	18.3	251	12	US-10-127-890-2	Sequence 2, App1
22	173.5	18.3	251	12	US-10-127-890-102	Sequence 102, App
23	173.5	18.3	251	12	US-10-127-890-103	Sequence 103, App
24	173.5	18.3	251	12	US-10-127-890-104	Sequence 104, App
25	173.5	18.3	251	12	US-10-127-890-105	Sequence 105, App
26	173.5	18.3	251	12	US-10-127-890-109	Sequence 109, App
27	172.5	18.2	293	9	US-09-765-527-259	Sequence 259, App
28	172.5	18.2	309	9	US-09-765-527-253	Sequence 253, App
29	172.5	18.2	332	9	US-09-765-527-251	Sequence 251, App
30	171.5	18.1	251	12	US-10-127-890-108	Sequence 108, App
31	125.5	13.2	314	10	US-09-978-274A-2	Sequence 2, App1
32	116.5	12.3	261	12	US-10-127-890-9	Sequence 9, App1
33	112.5	11.9	250	10	US-09-792-793A-36	Sequence 36, App1
34	112.5	11.9	250	12	US-10-127-890-8	Sequence 8, App1
35	111	11.7	263	10	US-09-978-274A-6	Sequence 6, App1
36	111	11.7	263	10	US-09-978-274A-4	Sequence 4, App1
37	93	9.8	280	12	US-10-127-890-3	Sequence 3, App1
38	93	9.8	280	12	US-10-127-890-127	Sequence 127, App
39	93	9.8	280	12	US-10-127-890-128	Sequence 128, App
40	93	9.8	280	12	US-10-127-890-129	Sequence 129, App
41	86.5	9.1	332	10	US-09-792-793A-73	Sequence 73, App1
42	86.5	8.9	332	10	US-09-792-793A-82	Sequence 82, App1
43	83.5	8.8	332	10	US-09-792-793A-76	Sequence 76, App1
44	79.5	8.4	254	10	US-09-792-793A-85	Sequence 85, App1
45	79.5	8.4	259	12	US-10-127-890-10	Sequence 10, App1

#### ALIGNMENTS

RESULT 1  
US-10-127-890-1  
; Sequence 1, Application US/10127890  
; Publication No. US20030166196A1  
GENERAL INFORMATION:

APPLICANT: Belter, Marc D.

Carroll, Stephen F.

Studnicka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-MAY-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-MAY-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-127-890-6

Query Match 28.0%; Score 265.5; DB 12; Length 247;  
Best Local Similarity 42.1%; Pred. No. 1.3e-21;  
Matches 56; Conservative 32; Mismatches 38; Indels 7; Gaps 4;

QY 44 INFTAGATVOSTYFNIRAVRGRLTTGADVREHPIVLPNVRGLPINQRFILVELSNHAEI 103  
DB 2 VSFRLSGATSSSYGVFISNLKALPNERKL-YDIPLL--RSSLPQSQRVALIHLTNVADE 58  
QY 104 SVTLADVTNAYVYVGRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVTFAGNGYDRLE 162  
DB 59 TISVAIDVTNAYVYIMGYRAGDTSYF--NEASATEAKYVFKDMRKVTLPLYSQNYERLQ 115  
QY 163 QLAGNIRENIELG 175  
DB 116 TAAGKIRENIEPLG 128

RESULT 5  
US-10-280-679B-4  
Sequence 4, Application US/10280679B  
Publication No. US20030150019A1  
GENERAL INFORMATION:  
APPLICANT: Large Scale Biology Corporation  
TITLE OF INVENTION: Monopartite RNA Virus Transformation Vectors  
FILE REFERENCE: LSBC-0109-US03  
CURRENT APPLICATION NUMBER: US/10/280,679B  
CURRENT FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 09/557,941  
PRIOR FILING DATE: 2000-04-24  
PRIOR APPLICATION NUMBER: 08/484,341  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: 07/923,692  
PRIOR FILING DATE: 1992-07-31  
PRIOR APPLICATION NUMBER: 07/600,244  
PRIOR FILING DATE: 1990-10-22  
PRIOR APPLICATION NUMBER: 07/641,617  
PRIOR FILING DATE: 1991-01-16  
PRIOR APPLICATION NUMBER: 07/737,899  
PRIOR FILING DATE: 1991-07-26  
PRIOR APPLICATION NUMBER: 07/739,143  
PRIOR FILING DATE: 1991-08-01  
PRIOR APPLICATION NUMBER: 07/310,881  
PRIOR FILING DATE: 1989-02-17

PRIOR APPLICATION NUMBER: 07/160,766  
PRIOR FILING DATE: 1988-02-26  
PRIOR APPLICATION NUMBER: 07/160,771  
PRIOR FILING DATE: 1988-02-26  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 289  
TYPE: PRT  
ORGANISM: Chinese cucumber protein alpha-trichosanthin  
US-10-280-679B-4

Query Match 28.0%; Score 265.5; DB 12; Length 289;  
Best Local Similarity 42.1%; Pred. No. 1.6e-21;  
Matches 56; Conservative 32; Mismatches 38; Indels 7; Gaps 4;

QY 44 INFTAGATVOSTYFNIRAVRGRLTTGADVREHPIVLPNVRGLPINQRFILVELSNHAEI 103  
DB 25 VSFRLSGATSSSYGVFISNLKALPNERKL-YDIPLL--RSSLPQSQRVALIHLTNVADE 81  
QY 104 SVTLADVTNAYVYVGRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRVTFAGNGYDRLE 162  
DB 82 TISVAIDVTNAYVYIMGYRAGDTSYF--NEASATEAKYVFKDMRKVTLPLYSQNYERLQ 138  
QY 163 QLAGNIRENIELG 175  
DB 139 TAAGKIRENIEPLG 151

RESULT 6  
US-10-282-935-3  
Sequence 3, Application US/10282935  
Publication No. US20030143193A1  
GENERAL INFORMATION:  
APPLICANT: VITETTA, ELLEN S.  
APPLICANT: GHETTE, VICTOR F.  
APPLICANT: SMALLSHAW, JOAN  
APPLICANT: BALUNA, ROXANA G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF  
FILE REFERENCE: UTSD:884US  
CURRENT APPLICATION NUMBER: US/10/282,935  
CURRENT FILING DATE: 2002-10-29  
PRIOR APPLICATION NUMBER: 09/538,873  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: 60/126,826  
PRIOR FILING DATE: 1999-03-30  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 251  
TYPE: PRT  
ORGANISM: Abtius precatorius  
US-10-282-935-3

Query Match 25.8%; Score 245; DB 12; Length 251;  
Best Local Similarity 44.8%; Pred. No. 2.6e-19;  
Matches 60; Conservative 18; Mismatches 46; Indels 10; Gaps 4;

QY 44 INFTAGATVOSTYFNIRAVRGRLTTGADVREHPIVLPNVRGLPINQRFILVELSNHAEI 103  
DB 5 IKFSTGATSSQSKYGEIHALRERLKG--LHIDIPVLPDPTTLQSRNRRITVELSNSTE 62  
QY 104 SVTLADVTNAYVYVGRAGNSAYFFH--PDNOEDAETHLFTDVQNRVTFAGNGYDRLE 161  
DB 63 SIEVGIDVTNAYVVAARACTOSYFLRDASSASD-----YLETGT-DQHSIPFGTYGIDL 116  
QY 162 EOLAGNIRENIELG 175  
DB 117 ERMAHOSKQRIPLG 130





Db 68 TAAIDVTNLYVAVYAGQDSYFLR-DAPRGAE--TILFTGT-TRSSLPFGNSYPLDERYA 123  
QY 166 GNIRENIEIG 175  
Db 124 GH-RDQIPIG 132

## RESULT 10

US-09-347-064-8  
Sequence 8, Application US/09347064A  
Patent No. US20020045208A1  
GENERAL INFORMATION:  
APPLICANT: Eck, Jurgen  
APPLICANT: Schmidt, Arno  
APPLICANT: Zinke, Holger  
TITLE OF INVENTION: Recombinant Fusion Proteins Based on  
TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum  
TITLE OF INVENTION: album  
FILE REFERENCE: 09282-5  
CURRENT APPLICATION NUMBER: US/09/347,064A  
CURRENT FILING DATE: 1999-07-02  
EARLIER APPLICATION NUMBER: PCT/EP98/00009  
EARLIER FILING DATE: 1998-01-02  
EARLIER APPLICATION NUMBER: EP 97 10 0012.0  
EARLIER FILING DATE: 1997-01-02  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 252  
TYPE: PRT  
ORGANISM: Viscum album  
US-09-347-064-8

## Query Match

Best Local Similarity 42.3%; Score 206.5; DB 9; Length 252;  
Matches 55; Conservative 24; Mismatches 42; Indels 9; Gaps 7;

QY 48 TAGATVOSTYNTFIRAVRGRLTTGADVRRHETPYLPNRVGLPIN--ORFLIVELSNHAEISV 105  
Db 9 TQTGTGEYFRITILRDYVSSGS--FSNEIPL--RQSTIPVSDAQRFVLVELTNGGDSI 66  
QY 106 TTAIDVTNLYVAVYAGQDSYFLR-DAPRGAE--TILFTGT-TRSSLPFGNSYPLDERYA 165  
Db 67 TAAIDVTNLYVAVYAGQDSYFLR-DAPRGAE--TILFTGT-TRSSLPFGNSYPLDERYA 122  
QY 166 GNIRENIEIG 175  
Db 123 GH-RDQIPIG 131

## RESULT 11

US-10-127-890-7  
Sequence 7, Application US/10127890  
Publication No. US20030166196A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-May-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-May-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-May-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-Dec-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-Jun-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-Nov-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248

## INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-10-127-890-7

Query Match 21.6%; Score 204.5; DB 12; Length 263;  
Best Local Similarity 37.3%; Pred. No. 1e-14;  
Matches 50; Conservative 30; Mismatches 45; Indels 9; Gaps 5;

QY 44 INFATCATVOSTYNTFIRAVRGRLTTGADVRRHETPYLPNRVGLPINORFLIVELSNHAE 102  
Db 2 VSFPLSGADPSRSTGMFKDLRNALPFRKXV-YNIPLLPSVSGA---GRYLAHRLNYDG 57  
QY 103 LSVTLADVTNLYVAVYAGQDSYFLR-DAPRGAE--TILFTGT-TRSSLPFGNSYPLDERYA 161  
Db 58 KTIIVADVTVNLYVAVYAGQDSYFLR-DAPRGAE--TILFTGT-TRSSLPFGNSYPLDERYA 122  
QY 162 GNIRENIEIG 175  
Db 115 QIAGKPRKIPIG 128

## RESULT 12

US-10-127-890-5  
Sequence 5, Application US/10127890  
Publication No. US20030166196A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07//87,567
FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 110:
SEQUENCE CHARACTERISTICS:
LENGTH: 251 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 110:
US-10-127-890-110

Query Match      18.7%; Score 177.5; DB 12; Length 251;
Best Local Similarity 37.6%; Pred. No. 1e-11; Indels 17; Gaps 8
Matches 53; Conservative 22; Mismatches 49;

OY      44      INFTTGATVQVOSTNFIKRAVRGLTLTGADVRIHLPVLRVGLPIHQRFILVELSNMHL 103
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       5      VSSITGATVTYTVNLTNLRYKLKPDEGN-SHGIPLLRKKADDP-GKAFVLVALSDNGC 62
          :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      104     SVTLALDYNNAVYVGIRAGNSAIFPH--PDNDQDAEATHLFTD-VQNRTFFPGGYDR 160
          ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       63      LAELAIIDVSVYVYGQVRNRSYFEFKDAP-----AAIEGLFKNTIKTR-LHRGGSYPS 115
          :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      161     LEOLAGN--LRENIELGNPL 179
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      116     LK---GEKARETTDIGIEPL 133
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-10-127-890-111
; Sequence 111, Application US/10127890
; Publication No. US2003016196A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
Carroll, Stephen F.
Studnicka, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mcandrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:

```

[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:40:32 ; Search time 11.6091 Seconds  
(without alignments)  
1482.817 Million cell updates/sec

Title: US-10-083-336a-2  
Perfect score: 948  
Sequence: 1 MKPGNTIYIMYAVATWLC.....RLQLAGNLRNIEIGNGPL 179

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR:76:\*  
2: PIR:1:\*  
3: PIR:3:\*  
4: PIR:4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	948	100.0	576	1	RLCSD
2	813.5	85.8	564	1	ricin D precursor
3	285	30.1	562	2	agglutinin precursor
4	282	29.7	528	2	abrin-c precursor
5	269	28.4	527	2	abrin-d precursor
6	265.5	28.0	289	1	RNA N-glycosidase
7	265	28.0	251	2	abrin (clone 7.2)
8	261.5	27.6	247	2	karasurin - Mongol
9	261.5	27.6	247	2	karasurin-B - Tric
10	261.5	27.6	247	2	karasurin C - Tric
11	245	25.8	528	1	abrin-a precursor
12	230	24.3	278	2	beta-luffin - smoo
13	218	23.0	250	2	luffin-B - smooth
14	216.5	22.8	254	2	misletoe lectin I
15	210.5	22.2	286	2	RNA N-glycosidase
16	208.5	22.0	286	2	RNA N-glycosidase
17	208	21.9	316	2	RNA N-glycosidase
18	207.5	21.9	245	2	RNA N-glycosidase
19	205.5	21.7	286	1	RNA N-glycosidase
20	195	20.6	277	2	RNA N-glycosidase
21	185.5	19.6	270	2	agglutinin I precu
22	136.5	14.4	294	2	RNA N-glycosidase
23	134	14.1	313	2	RNA N-glycosidase
24	118	12.4	278	2	RNA N-glycosidase
25	116.5	11.3	261	2	RNA N-glycosidase
26	110	11.6	289	2	antiviral protein
27	104.5	11.0	272	2	RNA N-glycosidase
28	103.5	10.9	278	2	betavulgin - beet
29	101	10.7	301	1	RNA N-glycosidase

30	99	10.4	293	2	S17519	RNA N-glycosidase
31	95	10.0	302	2	S11860	RNA N-glycosidase
32	95	10.0	304	2	S11859	RNA N-glycosidase
33	94	9.9	280	2	UC5848	protein synthesis
34	93	9.8	280	1	RLBH	RNA N-glycosidase
35	93	9.8	281	2	B38664	30K ribosome inact
36	92.5	9.8	492	2	S03098	jasmonate-induced
37	91.5	9.7	560	2	S25092	RNA N-glycosidase
38	90	9.5	292	1	RLQHG2	tritin - wheat
39	89	9.4	275	2	S33631	RNA N-glycosidase
40	89	9.4	283	2	S05205	aerolysin precursor
41	87.5	9.2	486	2	A25976	conserved hypotet
42	87.5	9.2	555	2	AC1203	conserved hypotet
43	87.5	9.2	557	2	A11560	conserved hypotet
44	85.5	9.0	157	2	S17934	RNA N-glycosidase
45	85.5	9.0	485	2	S26575	hemolysin - Aeromo

## ALIGNMENTS

### RESULT 1

RLCSD  
ricin D precursor - castor bean  
N:Contains: RNA N-glycosidase (EC 3.2.2.22)  
C:Species: Ricinus communis (castor bean)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
C:Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903  
R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.F.  
Nucleic Acids Res. 13, 8019-8033, 1985  
A>Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.  
A:Reference number: A24041; MUID:86067214; PMID:2999712  
A:Accession: A24041  
A:Molecule type: DNA  
A:Residues: 1-576 <HA>  
A:Cross-References: GB:X03179; NID:921082; PIDN:CAA26939.1; PID:921083  
R:Tregear, J.W.; Roberts, L.M.  
Plant Mol. Biol. 18, 515-525, 1992  
A>Title: The lectin gene family of Ricinus communis: cloning of a functional ricin ge  
A:Reference number: S20513; MUID:92163016; PMID:1371405  
A:Accession: S20513  
A:Molecule type: DNA  
A:Residues: 1-576 <TR>  
A:Cross-References: EMBL:X52908; NID:921084; PIDN:CAA37095.1; PID:921085  
R:Lamp, F.I.; Roberts, L.M.; Lord, J.M.  
Eur. J. Biochem. 148, 265-270, 1985  
A>Title: Nucleotide sequence of cloned cDNA coding for preproricin.  
A:Reference number: A24614; MUID:85179479; PMID:3838723  
A:Accession: A24614  
A:Molecule type: mRNA  
A:Residues: 12-75, 'D', '77-550, 'R', '552-576 <LAW>  
A:Cross-References: GB:X02388; NID:921077; PIDN:CAA26230.1; PID:921078  
R:Yoshitake, S.; Funatsu, G.; Funatsu, M.  
Agric. Biol. Chem. 42, 1267-1274, 1978  
A>Title: Isolation and sequences of peptic peptides, and the complete sequence of the  
A:Reference number: A03372  
A:Accession: A03372  
A:Molecule type: Protein  
A:Residues: 36-97, 'Q', '99-109, 'S', '111-269, 'D', '272-283, 'L', '285-288, 290-302 <YOS>  
A>Note: This paper cites the others in the series providing experimental details for  
R:Araki, T.; Funatsu, G.  
FEBS Lett. 191, 121-124, 1985  
A>Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptop  
A:Reference number: A24010  
A:Accession: A24010  
A:Molecule type: Protein  
A:Residues: 315-383, 'PS', '386-576 <ARA>  
R:Funatsu, G.; Kimura, M.; Funatsu, M.  
Agric. Biol. Chem. 43, 2221-2224, 1979  
A>Title: Primary structure of Ala chain of ricin D.  
A:Reference number: A03374  
A:Accession: A03374  
A:Molecule type: protein









Db 116 IAAKIRENIPLG 128

# RESULT 9

JC5032

Karasurin-B - Trichosanthes kirilowii var. japonica

C:Species: Trichosanthes kirilowii var. japonica

C:Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 23-May-1997

C:Accession: JC5032

R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.

Biol. Pharm. Bull. 19, 1485-1489, 1996

A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka

A:Reference number: JC5032; MUID:97108848; PMID:8951169

A:Accession: JC5032

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-247 <KON>

C:Comment: This protein belongs to type I ribosomal-inactivating proteins which catalyze

C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology

F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match

Best Local Similarity 42.98; Pred. No. 2e-16;

Matches 57; Conservative 32; Mismatches 37; Indels 7; Gaps 4;

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Db 116 IAAKIRENIPLG 128

Query 44 INFTAGATVQSTNTFRVAVRGRLLTGADVRHEIPVLPNVRGLPIPNORFIVELSNHAEI 103

Db 25 VSRFLSGATSSSYGVFISNLKRLPYERKL-YDIPLL--RSTLPGSQRFVALLHNTNYADE 81

Query 104 SVTLADVTNAYVYVGRAGNSAFEPHPDNOEDA-EAITHLEFDVQNRRTFAFGNDRLE 162

Db 82 TISVAIDVTNAYVYVGRAGNDSYF--NEASATEAKYVFKDKRKVTLPYSGNYERLQ 138

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175

Db 139 IAAKIRENIPLG 151

Query 163 QLAGNRENIETG 175







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:30:32 ; Search time 6.2752 Seconds  
(without alignments) 1341.437 Million cell updates/sec

Title: US-10-083-336a-2

Perfect score: 948  
Sequence: 1 MKPGNTIYIMYAVATWLC.....RLQLAGNLRENIELGNGPL 179

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Database : SwissProt\_41.\*  
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	948	100.0	576	1 RICI_RICCO	P02879 ricinus com
2	813.5	85.8	564	1 AGGL_RICCO	P06750 ricinus com
3	285	30.1	562	1 ABRG_ABRPR	P28590 abrus prec
4	269	28.4	527	1 ABRB_ABRPR	Q06077 abrus prec
5	265.5	28.0	289	1 RIPP_TRIKI	P09989 trichosan
6	261.5	27.6	282	1 RIPP_BRYDI	P98184 bryonia dio
7	261.5	27.6	289	1 RIPS_TRIKI	P24478 trichosan
8	245	25.8	528	1 ABRB_ABRPR	P11140 abrus prec
9	240.5	25.4	290	1 RIPP_BRYDI	P33185 bryonia dio
10	236	24.9	563	1 NIGB_SAMNI	P33183 sambucus nl
11	218	23.0	250	1 RIPP_LUPCY	P22851 luffa cylin
12	216.5	22.8	254	1 MVA_VISAL	P81446 viscum albu
13	210.5	22.2	286	1 RIPP_MOMBA	P29339 momordica b
14	209	22.0	286	1 RIPP_CUCFI	Q097X4 cucumis fig
15	208	21.9	316	1 RIPP_GELMU	P33186 gelonium mu
16	205.5	21.7	286	1 RIPP_MOMCH	P16094 momordica c
17	195	20.6	277	1 RIPP_LUPCY	Q00465 luffa cylin
18	189.5	20.0	294	1 RIPP_TRIAN	P56626 trichosan
19	136.5	14.4	334	1 RIPP_PHYAM	Q03464 physolacca
20	134	14.1	278	1 RIPP_PHYAM	P10297 physolacca
21	118	12.4	278	1 RIPP_MIRJA	P21326 mirabilis j
22	116.5	12.3	261	1 RIPS_PHYAM	P23339 physolacca
23	103	10.9	300	1 RIPP_MAIZE	P23591 zea mays (m
24	101	10.7	301	1 RIPP_MAIZE	P23593 zea mays (m
25	99	10.4	293	1 RIPP_MAIZE	P23591 zea mays (m
26	95	10.0	280	1 RIPP_MAIZE	P23592 zea mays (m
27	93	9.8	280	1 RIPP_HORVU	P22244 hordeum vul
28	93	9.8	280	1 RIPP_HORVU	P04399 hordeum vul
29	92.5	9.8	482	1 AERA_AERTR	P09166 aeromonas t
30	91.5	9.7	560	1 JI60_HORVU	Q00531 hordeum vul
31	90	9.5	292	1 RIPP_SAPOF	P27559 saponaria o
32	89	9.4	299	1 RIPP_SAPOF	P20656 saponaria o
33	85.5	9.0	157	1 RIPP_SAPOF	P27561 saponaria o

## ALIGNMENTS

RESULT 1	ID	RICI_RICCO	STANDARD:	PRT:	576 AA.
AC	P02879	RICI_RICCO	STANDARD:	PRT:	576 AA.
DT	21-JUL-1986	(Rel. 01, Created)			
DT	13-AUG-1987	(Rel. 05, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Ricin precursor [contains: Ricin A chain (rRNA N-glycosidase)				
DE	(EC 3.2.2.22): Ricin B chain).				
OS	Ricinus communis (Castor bean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosid 1; Malpighiales; Euphorbiaceae; Ricinus.				
OX	NCBI_TaxID=3988;				
ON	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=86067214; PubMed=2999712;				
RA	Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,				
RA	Weaver R.F.;				
RT	"Genomic cloning and characterization of a ricin gene from Ricinus				
RT	communis.";				
RL	Nucleic Acids Res. 13:8019-8033(1985).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92163016; PubMed=1371405;				
RA	Tregear J.W., Roberts L.M.;				
RT	"The lectin gene family of Ricinus communis: cloning of a functional				
RT	ricin gene and three lectin pseudogenes.";				
RL	Plant Mol. Biol. 18:515-525(1992).				
RN	[3]				
RP	SEQUENCE OF 12-576 FROM N.A.				
RX	MEDLINE=85179479; PubMed=3838723;				
RA	Lamb A., Roberts L.M., Lord J.M.;				
RT	"Nucleotide sequence of cloned cDNA coding for preproricin.";				
RL	Eur. J. Biochem. 148:265-270(1985).				
RN	[4]				
RP	SEQUENCE OF 36-302.				
RA	Yoshitake S., Funatsu G., Funatsu M.;				
RT	"Isolation and sequences of peptic peptides, and the complete				
RT	sequence of the chain of ricin-D.";				
RL	Agric. Biol. Chem. 42:1267-1274(1978).				
RN	[5]				
RP	SEQUENCE OF 315-576.				
RA	Funatsu G., Kimura M., Funatsu M.;				
RT	"Primary structure of Ala chain of ricin D.";				
RL	Agric. Biol. Chem. 43:2221-2224(1979).				
RN	[6]				
RP	CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.				
RX	MEDLINE=90344223; PubMed=1368517;				
RA	Kimura Y., Kusuhoku H., Tada M., Takagi S., Funatsu G.;				
RT	"Structural analyses of sugar chains from ricin A-chain variant.";				
RL	Agric. Biol. Chem. 54:157-162(1990).				
RN	[7]				
RP	REVIEW.				
RX	MEDLINE=21480122; PubMed=11595634;				
RA	Olines S., Kozlov J.V.;				

34	85.5	9.0	485	1 AER5_AERHY	006306 aeromonas h
35	85.5	9.0	493	1 AERA_AERHY	P09167 aeromonas h
36	83.5	8.8	253	1 RIPP_SAPOF	Q41389 saponaria o
37	83.5	8.8	253	1 RIPP_SAPOF	Q41391 saponaria o
38	82.5	8.7	262	1 VMEI_FIPV	P25878 felline infe
39	81	8.5	305	1 Y348_MYCPN	P75255 mycoplasma
40	80.5	8.5	492	1 AERA_AERHY	006303 aeromonas h
41	78.5	8.3	489	1 AERA_AERSA	008676 aeromonas s
42	78.5	8.3	492	1 AERA_AERHY	006305 aeromonas s
43	78.5	8.3	692	1 PHLN_PSEAE	P15713 pseudomonas
44	78	8.2	776	1 VP4_ROTET	Q02945 equine rota
45	76.5	8.1	488	1 AERA_AERSO	006304 aeromonas s

RT "Ricin.";  
 RL Toxicon 39:1723-1728(2001).  
 [8]  
 RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RP MEDLINE=97165983; PubMed=9558397;  
 RA Monfort W., Villafraña J.E., Monzingo A.F., Ernst S.R., Katzin B.,  
 RA Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.;  
 RT "The three-dimensional structure of ricin at 2.8 A.";  
 RL J. Biol. Chem. 262:5398-5403(1987).  
 [9]  
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
 RP MEDLINE=91352004; PubMed=1881881;  
 RA Katzin B.J., Collins E.J., Robertus J.D.;  
 RT "Structure of ricin A-chain at 2.5 A.";  
 RL Proteins 10:251-259(1991).  
 [10]  
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.  
 RP MEDLINE=91352005; PubMed=1881882;  
 RA Rutenber E., Robertus J.D.;  
 RT "Structure of ricin B-chain at 2.5-A resolution.";  
 RL Proteins 10:260-269(1991).  
 [11]  
 RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.  
 RP MEDLINE=95082010; PubMed=790130;  
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,  
 RA Pauplit R.A.;  
 RT "X-ray structure of recombinant ricin A-chain at 1.8-A resolution.";  
 RL J. Mol. Biol. 244:410-422(1994).  
 [12]  
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.  
 RP MEDLINE=96374222; PubMed=8780513;  
 RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,  
 RA Molina-Svint M.C., Robertus J.D.;  
 RT "Structure and activity of an active site substitution of ricin A chain.";  
 RL Biochemistry 35:11098-11103(1996).  
 [13]  
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
 RP MEDLINE=97240820; PubMed=9086280;  
 RA Yan X., Hollis T., Svint M., Day P., Monzingo A.F., Milne G.W.,  
 RA Robertus J.D.;  
 RT "Structure-based identification of a ricin inhibitor.";  
 RL J. Mol. Biol. 266:1043-1049(1997).  
 [14]  
 RN MUTAGENESIS.  
 RP MEDLINE=93165632; PubMed=1287657;  
 RA Kin Y., Robertus J.D.;  
 RT "Analysis of several key active site residues of ricin A chain by mutagenesis and x-ray crystallography.";  
 RL Protein Eng. 5:775-779(1992).  
 -i- FUNCTION: Ricin is highly toxic to animal cells and to a less extent to plant cells. The A chain is responsible for inhibiting protein synthesis through the catalytic inactivation of 60S ribosomal subunits. It acts as a glycosidase that removes a specific adenine residue from an exposed loop of 28S ribosomal RNA. As this loop is involved in the binding of elongation factors, the modified ribosomes are unable to support protein synthesis. The A chain can inactivate a few thousand ribosomes per minute, thus inactivating them faster than the cell can make new ones. A single A-chain molecule can therefore kill an animal cell. The B chain binds to cell receptors and facilitates the entry into the cell of the A chain; B chains are also responsible for cell agglutination (Lectin activity). It binds to beta-D-galactopyranoside moieties.  
 -i- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 -i- SUBUNIT: Disulfide-linked dimer of A and B chains.  
 -i- DOMAIN: The B chain is composed of two domains, each domain consists of 3 homologous subdomains (alpha, beta, gamma).  
 -i- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.  
 -i- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.

CC -i- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -i- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS  
 CC WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, AND REF.3).  
 CC -i- DATABASE: NAME-Protein Spotlights;  
 CC NOTE=Issue 31 of February 2003;  
 CC WWW="http://www.expasy.org/spotlight/articles/splct031.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; X03179; CAA26939.1; -;  
 DR EMBL; X52908; CAA37095.1; -;  
 DR EMBL; X02388; CAA26230.1; -;  
 DR EMBL; A12892; CAA01058.1; -;  
 DR PIR; A24041; RLCSO.  
 DR PDB; 2AA1; 31-JAN-94.  
 DR PDB; 1APG; 31-JAN-94.  
 DR PDB; 1FMP; 31-OCT-93.  
 DR PDB; 1IFS; 14-JAN-98.  
 DR PDB; 1IFU; 14-JAN-98.  
 DR PDB; 1IFC; 31-OCT-93.  
 DR PDB; 1OBS; 16-JUN-97.  
 DR PDB; 1OBT; 16-JUN-97.  
 DR PDB; 1BR5; 02-SEP-98.  
 DR PDB; 1BR6; 02-SEP-98.  
 DR PDB; 1IL3; 16-JAN-02.  
 DR PDB; 1IL4; 16-JAN-02.  
 DR PDB; 1IL9; 16-JAN-02.  
 DR GLYCOSULEDB; P02879; -;  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 DR Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; Signal; 3D-structure.  
 KW SIGNAL 1 35  
 FT CHAIN 36 302  
 FT PEPTIDE 303 314  
 FT CHAIN 315 576  
 FT DOMAIN 321 448  
 FT DOMAIN 451 575  
 FT REPEAT 331 373  
 FT REPEAT 374 414  
 FT REPEAT 417 449  
 FT REPEAT 462 497  
 FT REPEAT 501 540  
 FT REPEAT 543 570  
 FT ACT\_SITE 212 212  
 FT DISULFID 294 318  
 FT DISULFID 334 353  
 FT DISULFID 377 394  
 FT DISULFID 465 478  
 FT DISULFID 504 521  
 FT CARBOHYD 45 45  
 FT CARBOHYD 271 271  
 FT CARBOHYD 409 409  
 FT CARBOHYD 449 449  
 FT CONFLICT 76 76  
 FT CONFLICT 551 551  
 FT STRAND 43 47  
 FT TURN 49 50  
 N-LINKED (GLCNAC. . .).  
 /FTID-CAR.000080.  
 N-LINKED (GLCNAC. . .) (IN MINOR FORM).  
 /FTID-CAR.000081.  
 N-LINKED (GLCNAC. . .).  
 N-LINKED (GLCNAC. . .).  
 E -> D (IN REF. 3).  
 A -> R (IN REF. 3).  
 INTERCHAIN.

Query Match 100.0%; Score 948; DB 1; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 1,4e-83;  
 Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPGGTYIWMYAVATWLCFGSTSGMSFTLEDDNIPPKOYPIINFTTAGATVOSYTNFI 60  
 |||||||  
 DB 1 MKPGGTYIWMYAVATWLCFGSTSGMSFTLEDDNIPPKOYPIINFTTAGATVOSYTNFI 60  
 |||||||

OY 61 RAVRGRLTGADVRHEIPVLPNRVGLPINORFIVELSNHAEISVTLADVTNAYVGYR 120  
 |||||||  
 DB 61 RAVRGRLTGADVRHEIPVLPNRVGLPINORFIVELSNHAEISVTLADVTNAYVGYR 120  
 |||||||

OY 121 AGNSAYFFHPNDQEDAEATITLFTDVQNRRTFAFGNDRLEQLAGNLRENIELGNGPL 179  
 |||||||  
 DB 121 AGNSAYFFHPNDQEDAEATITLFTDVQNRRTFAFGNDRLEQLAGNLRENIELGNGPL 179  
 |||||||

RESULT 2  
 AGSL\_RICCO STANDARD; PRT; 564 AA.  
 ID AGSL\_RICCO  
 AC P06750;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (rRNA N-  
 DE glycosidase) (EC 3.2.2.22); Agglutinin B chain].  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
 OX NCBI\_TaxID=3986;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86059449; PubMed=2999130;  
 RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;  
 RT "The primary sequence of Ricinus communis agglutinin. Comparison with  
 RL ricin.",  
 RL J. Biol. Chem. 260:15682-15686(1985).  
 RN [2]  
 RP SEQUENCE OF 303-564.  
 RC Tissue=Seed;  
 RA Araki T., Yoshida Y., Funatsu G.;  
 RT "The complete amino acid sequence of the B-chain of the Ricinus  
 RL communis agglutinin isolated from large-grain castor bean seeds.",  
 RL Biochim. Biophys. Acta 872:277-285(1986).  
 RN [3]  
 RP SEQUENCE OF 303-337.  
 RX MEDLINE=8018723; PubMed=6768555;  
 RA Lin T.T.-S., Li S.S.-L.;  
 RT "Purification and physicochemical properties of ricins and  
 RL agglutinins from Ricinus communis.",  
 RL Eur. J. Biochem. 105:453-459(1980).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: M12089; AAA33869.1; -;  
 DR EMBL: S40368; AB22584.1; -;  
 DR PIR: A24361; RLC5AC;  
 DR HSP: P02879; IIR6;  
 DR GlycoSuiteDB: P06750; -;  
 DR InterPro: IPR000772; Ricin\_B\_lectin.

DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS00231; RICIN\_B\_LLECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 290  
 FT PROPEP 291 302  
 FT CHAIN 303 364  
 FT DOMAIN 309 436  
 FT DOMAIN 439 563  
 FT REPEAT 319 361  
 FT REPEAT 362 402  
 FT REPEAT 405 437  
 FT REPEAT 450 485  
 FT REPEAT 489 528  
 FT REPEAT 531 558  
 FT ACT\_SITE 200 200  
 FT DISULFID 282 306  
 FT DISULFID 322 341  
 FT DISULFID 365 382  
 FT DISULFID 453 466  
 FT DISULFID 492 509  
 FT CARBOHYD 34 34  
 FT CARBOHYD 259 259  
 FT CARBOHYD 397 397  
 FT CARBOHYD 437 437  
 FT CONFLICT 331 331  
 FT CONFLICT 362 362  
 FT CONFLICT 374 374  
 FT CONFLICT 404 404  
 FT CONFLICT 552 552  
 SQ SEQUENCE 564 AA; 62851 MW; D455F2A72FE09759 CRC64;

Query Match 85.8%; Score 813.5; DB 1; Length 564;  
 Best Local Similarity 93.5%; Pred. No. 1,2e-70;  
 Matches 157; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

OY 12 MYAVATWLCFGSTSGMSFTLEDDNIPPKOYPIINFTTAGATVOSYTNFI RAVRGRLTGA 71  
 |||||||  
 DB 1 MYAVATWLCFGSTSGMSFTLEDDNIPPKOYPIINFTTAGATVOSYTNFI RAVRGRLTGA 60  
 |||||||

OY 72 DVHREIPVLPNRVGLPINORFIVELSNHAEISVTLADVTNAYVGYRAGNSAYFFHPD 131  
 |||||||  
 DB 61 DVHREIPVLPNRVGLPINORFIVELSNHAEISVTLADVTNAYVGYRAGNSAYFFHPD 120  
 |||||||

OY 132 NOEDAEATITLFTDVQNRRTFAFGNDRLEQLAGNLRENIELGNGPL 179  
 |||||||  
 DB 121 NOEDAEATITLFTDVQNRRTFAFGNDRLEQLAGNLRENIELGNGPL 167  
 |||||||

RESULT 3  
 ABRC\_ABRPR STANDARD; PRT; 562 AA.  
 ID ABRC\_ABRPR  
 AC P28590;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Abrin-c precursor [Contains: Abrin-c A chain (rRNA N-glycosidase)  
 DE (EC 3.2.2.22); Abrin-c B chain].  
 OS Abrus precatorius (Indian jicorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreace; Abrus.  
 OX NCBI\_TaxID=3816;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Leaf;  
 RX MEDLINE=91266957; PubMed=2050149;

RA Wood K.A., Lord J.M., Mawrzyniczak E.J., Platak M.;  
 RT "Preproabrin: genomic cloning, characterisation and the expression of  
 the A-chain in *Escherichia coli*.";  
 RL Eur. J. Biochem. 198;723-732(1991).  
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
 CC B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE  
 CC BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X55667; CAA39202.1; -.  
 CC PIR: S16022; S16022.  
 CC HSSP: P11140; IABR.  
 CC InterPro: IPR000772; Ricin\_B\_lectin.  
 CC InterPro: IPR001574; RIP.  
 CC Pfam: PF00652; Ricin\_B\_lectin; 6.  
 CC Pfam: PF00161; RIP; 1.  
 CC PRINTS: PR00396; SHIGARICIN.  
 CC SMART: SM00458; RICIN; 2.  
 CC DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 CC DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 CC KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 CC Glycoprotein; Lectin; Signal; Pyrrolidone carboxylic acid.  
 CC FT SIGNAL 1 34  
 CC FT CHAIN 1 285 ABRIN C A CHAIN (BY SIMILARITY).  
 CC FT PEPTIDE 286 295 LINKER PEPTIDE (BY SIMILARITY).  
 CC FT CHAIN 296 562 ABRIN C B CHAIN (BY SIMILARITY).  
 CC FT DOMAIN 307 434 RICIN B-TYPE LECTIN 1.  
 CC FT DOMAIN 437 561 RICIN B-TYPE LECTIN 2.  
 CC FT REPEAT 317 359 1-ALPHA.  
 CC FT REPEAT 360 400 1-BETA.  
 CC FT REPEAT 403 435 1-GAMMA.  
 CC FT REPEAT 448 483 2-ALPHA.  
 CC FT REPEAT 487 526 2-BETA.  
 CC FT REPEAT 529 562 2-GAMMA.  
 CC FT ACT\_SITE 198 198 INTERCHAIN (BY SIMILARITY).  
 CC FT DISULFID 281 303 BY SIMILARITY.  
 CC FT DISULFID 320 339 BY SIMILARITY.  
 CC FT DISULFID 363 380 BY SIMILARITY.  
 CC FT DISULFID 451 464 BY SIMILARITY.  
 CC FT DISULFID 490 507 BY SIMILARITY.  
 CC FT MOD\_RES 35 35 PYRROLIDONE CARBOXYLIC ACID  
 CC (BY SIMILARITY).  
 CC FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SEQUENCE 562 AA; 1JF0A8C7D7BA6278 CRC64;  
 CC -----

Query Match 30.1%; Score 285; DB 1; Length 562;  
 Best Local Similarity 44.1%; Pred. No. 8.6e-20;  
 Matches 71; Conservative 18; Mismatches 62; Indels 10; Gaps 4;

OY 19 ICFSTGSMFTLENNITP-----KQYPIINFTTAGAVSYTNFIRAVRGLTTGADV 74  
 DB 10 ICLATTCFSFALRCARARYPPVATQDVYKFTTGATGASQSYKQFETALRQLRG--LI 67  
 OY 75 HEIPVLPNRVGLPINORFLVELSNHAEISVTALDVTNATVYVGRAGNSAIFRHPNOE 134

DB 68 HDIPLPPTTVEERNKRIITVELNSERESIEVDITNATVAVRAGSQSYFL---RDA 124  
 OY 135 DAVALTHLETDVQNRFFAFGNYDRLEQLAGNIRENIEG 175  
 DB 125 PASASTYLEPGTQ-RYSLSRFDGSGYDLERMAHQREISISG 164  
 RESULT 4  
 ABRB\_ABRPR STANDARD; PRT; 527 AA.  
 AC 006077; P81374;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ABrin-b precursor [contains: ABrin-b A chain (rRNA N-glycosidase)  
 DE (EC 3.2.2.22); ABrin-b B chain].  
 OS Abrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosoid II; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 OX NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93132798; PubMed=8421313;  
 RA Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;  
 RT "Primary structure of three distinct isoabirins determined by cDNA  
 RT sequencing. Conservation and significance.";  
 RL J. Mol. Biol. 229:263-267(1993).  
 RN [2]  
 RP SEQUENCE OF 260-527.  
 RC TISSUE=Seed.  
 RX MEDLINE=93169023; PubMed=7763422;  
 RA Kimura M., Sumizawa T., Funatsu G.;  
 RT "The complete amino acid sequences of the B-chains of abrin-a and  
 RT abrin-b, toxic proteins from the seeds of Abrus precatorius.";  
 RL Biochim. Biophys. Acta 57:166-169(1993).  
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
 CC ABRIN-A IS MORE TOXIC THAN RICIN.  
 CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
 CC PRECEDES ENDOCYTOSIS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M98345; AAA3625.1; -.  
 CC PIR: S32430; S32430.  
 CC HSSP: P11140; IABR.  
 CC InterPro: IPR000772; Ricin\_B\_lectin.  
 CC InterPro: IPR001574; RIP.  
 CC Pfam: PF00652; Ricin\_B\_lectin; 6.  
 CC Pfam: PF00161; RIP; 1.  
 CC PRINTS: PR00396; SHIGARICIN.  
 CC SMART: SM00458; RICIN; 2.  
 CC DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 CC DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 CC KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;



KM Glycoprotein: Lectin; Pyrrolidone carboxylic acid.  
 FT CHAIN 1 250  
 FT PEPTIDE 251 260  
 FT CHAIN 261 527  
 FT DOMAIN 272 399  
 FT DOMAIN 402 526  
 FT REPEAT 282 324  
 FT REPEAT 325 365  
 FT REPEAT 368 400  
 FT REPEAT 413 448  
 FT REPEAT 452 491  
 FT REPEAT 494 527  
 FT ACT SITE 163 163  
 FT DISULFID 246 268  
 FT DISULFID 285 304  
 FT DISULFID 328 345  
 FT DISULFID 416 429  
 FT DISULFID 455 472  
 FT MOD\_RES 1 1  
 FT  
 FT CARBOHYD 110 110  
 FT CARBOHYD 360 360  
 FT CARBOHYD 400 400  
 FT CARBOHYD 282 282  
 FT CONFLICT 291 291  
 FT CONFLICT 350 351  
 FT CONFLICT 378 378  
 FT CONFLICT 426 426  
 FT CONFLICT 428 428  
 FT CONFLICT 431 431  
 FT CONFLICT 484 484  
 FT CONFLICT 491 491  
 FT CONFLICT 493 493  
 FT CONFLICT 502 502  
 FT CONFLICT 509 509  
 FT CONFLICT 513 513  
 FT CONFLICT 516 516  
 FT SEQUENCE 527 AA; 59114 MM; 3253AE490CE9494A CRC64;  
 Query Match Best Local Similarity 28.4%; Score 269; DB 1; Length 527;  
 Matches 65; Conservative 15; Mismatches 50; Indels 6; Gaps 3;  
 Oy 40 QYPIINFTAGATVQSYTNFIRAVRGRLTGGADVRHRIPLPVRVGLPINOPIFLVELSN 99  
 Db 1 QDQYIKFTTGATGTSQSKOTIEALRQRLTGG--LHGIPVLPDPPTLQERNRRTISVELSN 58  
 Oy 100 HAEISVTLADVTNAVYVGRAGNSAVFEHPDQEDAEATITLFTDVQRRTFAFGAND 159  
 Db 59 SDTESIAGIDVSNAYVAVRAGNRSYFL--RDAPTSASRYLFTGTQ-QYSLRFNGSYI 114  
 Oy 160 RLEQLAGNLRNENELG 175  
 Db 115 DLERLARQTRQOIPLG 130  
 RESULT 5  
 ID RIPT\_TRIKI STANDARD; PRT; 289 AA.  
 AC P09989;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DE 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ribosome-inactivating protein alpha-trichosanthin precursor  
 DE (tRNA N-glycosidase) (EC 3.2.2.22) (Alpha-TGS).  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).  
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;  
 OC Spermatophyta: Magnoliopsida: eudicotyledons: core eudicots: Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 ON NCBI\_TaxID=3677;  
 RX NCBI  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Maximowicz;

RX MEDLINE=91153657; PubMed=1999291;  
 RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;  
 RT "Cloning of trichosanthin cDNA and its expression in Escherichia  
 RT coli.";  
 RL Gene 97:267-272(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Maximowicz; TISSUE=leaf;  
 RX MEDLINE=90256790; PubMed=2341400;  
 RA Chow T., Feldman R.A., Lovett M., Piatak M.;  
 RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a  
 RT type I ribosome-inactivating protein.";  
 RL J. Biol. Chem. 265:8670-8674(1990).  
 RN [3]  
 RP SEQUENCE OF 24-270.  
 RC STRAIN=Maximowicz; TISSUE=Tuberos root;  
 RX MEDLINE=90256789; PubMed=2341399;  
 RA Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,  
 RA Wu P., Hwang K., Piatak M.;  
 RT "Primary amino acid sequence of alpha-trichosanthin and molecular  
 RT models for abrin A-chain and alpha-trichosanthin.";  
 RL J. Biol. Chem. 265:8665-8669(1990).  
 RN [4]  
 RP SEQUENCE OF 24-270.  
 RC TISSUE=Tuberos root;  
 RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,  
 RA Tian G.Y., Ni C.Z.;  
 RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and  
 RT application.";  
 RL Pure Appl. Chem. 58:789-798(1986).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).  
 RX MEDLINE=94344957; PubMed=806085;  
 RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;  
 RT "Structure of trichosanthin at 1.88-A resolution.";  
 RL Proteins 19:4-13(1994).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
 RX MEDLINE=95344383; PubMed=7619070;  
 RA Huang O., Liu S., Tang Y., Jin S., Wang Y.;  
 RT "Studies on crystal structures, active-centre geometry and  
 RT deproteinizing mechanism of two ribosome-inactivating proteins.";  
 RL Biochem. J. 309:285-298(1995).  
 CC CC -1- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS  
 CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT  
 CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.  
 CC CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC type I RIP SUBFAMILY.  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL: M34858; AAA34207.1; -;  
 CC EMBL: J05434; AAA34206.1; -;  
 CC PIR: J0566; RL727.  
 CC DR PDB: IMR; 07-FEB-95.  
 CC DR PDB: IMR; 07-FEB-95.  
 CC DR PDB: IMR; 10-JUL-95.  
 CC DR PDB: J34G; 28-JAN-03.  
 CC DR PDB: INI; 21-JAN-03.  
 CC DR PDB: IOD2; 24-APR-00.  
 CC DR InterPro: IPR001574; RIP.  
 CC DR Pfam: PF00161; RIP.  
 CC DR PRINTS: PR00396; SHIGARICIN.  
 CC DR PROSITE: PS00275; SHIGA\_RICIN.  
 CC Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;

KM	Toxin:	Signal;	3D-structure.
FT	SIGNAL	1	23
FT	CHAIN	24	270
FT	PROPER	271	289
FT	ACT_SITE	183	183
FT	CONFLICT	57	60
FT	CONFLICT	82	84
FT	CONFLICT	87	87
FT	CONFLICT	92	92
FT	CONFLICT	143	144
FT	CONFLICT	196	196
FT	CONFLICT	215	216
FT	CONFLICT	231	231
FT	CONFLICT	234	234
FT	CONFLICT	246	266
FT	CONFLICT	247	247
FT	STRAND	25	28
FT	TURN	30	31
FT	HELIX	34	46
FT	TURN	47	47
FT	STRAND	50	54
FT	TURN	55	56
FT	STRAND	66	69
FT	HELIX	70	76
FT	TURN	78	79
FT	STRAND	82	88
FT	TURN	89	92
FT	STRAND	93	99
FT	TURN	100	101
FT	STRAND	102	105
FT	HELIX	109	114
FT	TURN	115	117
FT	TURN	120	121
FT	STRAND	124	127
FT	HELIX	134	141
FT	TURN	142	142
FT	STRAND	145	147
FT	HELIX	150	150
FT	STRAND	152	163
FT	HELIX	164	165
FT	TURN	167	180
FT	HELIX	181	181
FT	TURN	182	186
FT	HELIX	187	187
FT	STRAND	188	195
FT	HELIX	196	196
FT	TURN	202	202
FT	STRAND	206	226
FT	HELIX	227	230
FT	TURN	231	239
FT	STRAND	241	242
FT	TURN	245	250
FT	HELIX	251	252
FT	TURN	254	258
FT	HELIX	259	259
FT	TURN	260	260
FT	STRAND	263	263
FT	TURN	266	268
STQ	SEQUENCE	289 AA; 31676 MM; 5CE09BB630575EB9 CRC64;	
<hr/>			
Query Match            28.0%; Score 265.5; DB 1: Length 289;			
Best Local Similarity 42.1%; Pred. No. 2.9e-18;			
Matches	Conservative	Matches	Mismatches
56;	32;	38;	Indels 7; Gaps 4;
OY	44	INFTAGAVOSTNFIKRVGRLLTGADVRHPIVLPNRVGLEPTIQRFIVELSNHAFL	103
DG	25	VSRILSGATSSSISGVGFISNLRKALPNRKL-YDIPL--RSSIPGSQRVALIHLTYADE	81
OY	104	SVTLADYNNAVYVYGRRAGSAVEFHPDDNOEDA-EAITHLFDPVCNRRTFAFGANYDRLE	162
DG	82	TTSVAIDVNVMYTMRAGDTSYFF--NEASATEAKYYFDAMARKVTLPYSNGNEREQ	138

Qy	163	QIAGNIRENIELG	175
Db	139	TAAGKIRENIPDG	151
RESULT 6			
RIP2_BRYDI	ID	RIP2_BRYDI	STANDARD; PRT; 282 AA.
AC	P98184;	Q9S8J0;	
DT	16-OCT-2001	(Rel. 40, Created)	
DT	16-OCT-2001	(Rel. 40, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last annotation update)	
DE	ribosome-inactivating protein bryodin II precursor (rRNA N-glycosidase) (EC 3.2.2.22) (BD2).		
OS	Bryonia dioica (Plant bryony).		
OC	Eukaryota; Viridipladetes; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucosids I; Cucurbitales; Cucurbitaceae; Bryonia.		
OC	NCBI_TaxID=3652;		
RP	[1]		
RA	SEQUENCE FROM N.A.		
RA	Stegall C.B., Gawlak S.L., Marguardt H.;		
RT	"Bryodin 2 a ribosome-inactivating protein isolated from the plant Bryonia dioica."		
RL	Patent number US5597569, 28-JAN-1997.		
RN	[2]		
RP	SEQUENCE OF 22-42.		
RC	TISSUE-ROOT;		
RX	MEDLINE=95151812; PubMed=7849072;		
RA	Stegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B., Marguardt H.;		
RT	"Characterization of ribosome-inactivating proteins isolated from Bryonia dioica and their utility as carcinoma-reactive immunoconjugates."		
RL	Bioconjug. Chem. 5:423-429(1994).		
CC	-I- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).		
CC	-I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.		
CC	-I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 1 RIP SUBFAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
CC	EMBL: T34238; -; NOT_ANNOTATED_CDS.		
DR	HSSP; P09989; IMR0.		
DR	InterPro; IPR001574; RIP.		
DR	Pfam; PF00161; RIP; 1.		
DR	PRINTS; PR00396; SHIGARICIN.		
DR	PROSITE; PS00275; SHIGA_RICIN; 1.		
KW	Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;		
KW	Multigene family; Glycoprotein; Signal.		
FT	STGNAL	1	21
FT	CHAIN	22	282
FT	ACT_SITE	183	
FT	CARBOHYD	25	25
SO	SEQUENCE	282 AA;	30754 MW; C52BE2P6A873769C CRC64;
Query Match 27.6%; Score 261.5; DB 1; Length 282;			
Best Local Similarity 50.0%; Pred. No. 6; 8e-18;			
Matches 67; Conservative 13; Mismatches 45; Indels 9; Gaps 5			
Qy	44	INFTTGATGVSTNFTIRAVGRILTGAVVRHIDVLPN-RVGLPINRFTLVELSNHAE	102
Db	24	INSLIGATGATKTTIRNLRTKLYGTGRVYVDIPVLRNAAAGL--ARFOLVITLTNYNG	80

QY 103 LSTTALDVTNAVVGVRAGNSAYFFHPDNOEDAETHFTDVOYRRTFAFGNDRL 162  
 DB 81 ESYTVLADVTNAVVGVRAGNTAFYFL---ADASTEANNVLFAGI-NHVRPLPGYNGDGL 136  
 OY 163 QLAGNL-RENIELG 175  
 DB 137 TAAGRISRENIELG 150

RESULT 7  
 RIPS\_TRIKI  
 ID RIPS\_TRIKI STANDARD: PRT: 289 AA.  
 AC P24478;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein karasurin precursor (rRNA N-glycosidase) (EC 3.2.2.22).  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 NCBI\_TaxID=3677;  
 RX MEDLINE-92005921; PubMed-1914000;  
 RA Toyokawa S., Takeda T., Kato Y., Wakabayashi K., Ogihara Y.;  
 RT "The complete amino acid sequence of an abortifacient protein, karasurin.";  
 RT Karasurin.";  
 RT Chem. Pharm. Bull. 39:1244-1249(1991).  
 CC -1- FUNCTION: ABORTION-INDUCING PROTEIN. IT INACTIVATES EUKARYOTIC  
 CC 60S RIBOSOMAL SUBUNITS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: AB000666; BAA21786.1; -.  
 DR PIR: JC3606; JC3606.  
 DR PIR: J00393; J00393.  
 DR HSSP: P09989; IMRJ.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP.1  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN.1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin; Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 270 KARASURIN-C.  
 FT PROPEP 271 289 REMOVED IN MATURE FORM.  
 FT ACT\_SITE 183 183 BY SIMILARITY.  
 SQ SEQUENCE 289 AA; 31704 MW; 883D3E3242887B26 CRC64;

Query Match 27.6%; Score 261.5; DB 1; Length 289;  
 Best Local Similarity 42.9%; Pred. No. 7e-18;

Matches 57; Conservative 32; Mismatches 37; Indels 7; Gaps 4;  
 OY 44 INFTAGATVOSTYTFEIRAVGRGLTTGADVHEIPVLNRYGLPINORFILVELSNHAE 103  
 DB 25 VSRFLSGATSSSYGVFINLKRALPYERKL-YDIPDL--RSTLFGSQRYALIHILTNAD 81  
 OY 104 SVTALDVTNAVVGVRAGNSAYFFHPDNOEDA-ETHFTDVOYRRTFAFGNDRL 162  
 DB 82 TISVALDVTNAVVGVRAGDTSYFF---NEASATEAKYFVKDAKKRYTLPSGNYERLQ 138

QY 163 QLAGNL-RENIELG 175  
 DB 139 IAAKRIRENIPLG 151

RESULT 8  
 ABRA\_ABRPR  
 ID ABRA\_ABRPR STANDARD: PRT: 528 AA.  
 AC P11140; P28589;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Abrin-a precursor [contains: Abrin-a A chain (rRNA N-glycosidase) (EC 3.2.2.22); Abrin-a B chain].  
 OS Abrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabiales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 NCBI\_TaxID=3816;  
 RX MEDLINE-93132798; PubMed-8421313;  
 RA Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;  
 RT "Primary structure of three distinct isoabris determined by cDNA  
 RT sequencing. Conservation and significance.";  
 RT J. Mol. Biol. 229:263-267(1993).  
 RN [12]  
 RN SEQUENCE OF 1-251.  
 RC TISSUE=Seed;  
 RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;  
 RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic  
 RT protein from the seeds of Abrus precatorius.";  
 RT Agric. Biol. Chem. 52:1095-1097(1988).  
 RN [13]  
 RN SEQUENCE OF 1-251 FROM N.A.  
 RC TISSUE=Leaf;  
 RX MEDLINE-91201329; PubMed-2016300;  
 RA Eversen G., Mathiesen A., Sundan A.;  
 RT "Direct molecular cloning and expression of two distinct abrin  
 RT A-chains.";  
 RT J. Biol. Chem. 266:6848-6852(1991).  
 RN [14]  
 RN SEQUENCE OF 262-528.  
 RX MEDLINE-92371656; PubMed-1505674;  
 RA Chen Y.-L., Chow L.-P., Tsungita A., Lin J.-Y.;  
 RT "The complete primary structure of abrin-a B chain.";  
 RT FEBS Lett. 309:115-118(1992).  
 RN [15]  
 RN X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).  
 RX MEDLINE-9533188; PubMed-7608980;  
 RA Tshilov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;  
 RT "Crystal structure of abrin-a at 2.14 A.";  
 RT J. Mol. Biol. 250:354-367(1995).  
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
 CC ABRIN-A IS MORE TOXIC THAN RICIN.  
 CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
 CC PRECEDES ENDOCYTOSIS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.

CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-  
CC INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.  
CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: M98344; AAA32624.1; ALT. INT.  
CC EMBL: X54872; -; NOT\_ANNOTATED\_CDS.  
CC PIR: S32429; TZLSN.  
CC PDB: 1ABR; 07-FEB-95.  
CC DR InterPro: IPR000772; Ricin\_B\_lectin.  
CC DR InterPro: IPR001574; RIP.  
CC DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
CC DR Pfam: PF00161; RIP; 1.  
CC DR PRINTS: PR00396; SHIGARICIN.  
CC DR SMART: SM00458; RICIN; 2.  
CC DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
CC DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
CC KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
CC KW Glycoprotein; Lectin; 3D-structure; Pyrrolidone carboxylic acid.  
CC CHAIN 1 251 ABRIN-A A CHAIN.  
CC DEPTIDE 252 261 LINKER PEPTIDE.  
CC CHAIN 262 528 ABRIN-A B CHAIN.  
CC DOMAIN 273 400 RICIN-B-TYPE LECTIN 1.  
CC DOMAIN 403 527 RICIN-B-TYPE LECTIN 2.  
CC REPEAT 283 325 1-ALPHA.  
CC REPEAT 326 366 1-BETA.  
CC REPEAT 369 401 1-GAMMA.  
CC REPEAT 414 449 2-ALPHA.  
CC REPEAT 453 492 2-BETA.  
CC REPEAT 495 528 2-GAMMA.  
CC ACT\_SITE 164 164 BY SIMILARITY.  
CC DISULFID 247 269 INTERCHAIN (BY SIMILARITY).  
CC DISULFID 286 305 BY SIMILARITY.  
CC DISULFID 329 346 BY SIMILARITY.  
CC DISULFID 417 430 BY SIMILARITY.  
CC DISULFID 456 473 BY SIMILARITY.  
CC MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
CC CARBOHYD 361 361 N-LINKED (GLCNAC. . .).  
CC CARBOHYD 401 401 MISSING (IN REF. 2).  
CC CONFLICT 202 202 N-> Y (IN REF. 4).  
CC CONFLICT 298 298 M-> L (IN REF. 4).  
CC CONFLICT 427 427 T-> P (IN REF. 4).  
CC CONFLICT 467 467 V-> L (IN REF. 4).  
CC CONFLICT 483 483  
CC STRAND 5 8  
CC TURN 10 11  
CC HELIX 14 28  
CC STRAND 32 33  
CC TURN 34 35  
CC STRAND 36 38  
CC TURN 42 43  
CC HELIX 47 49  
CC STRAND 51 57  
CC STRAND 63 69  
CC TURN 70 72  
CC STRAND 75 79  
CC STRAND 83 86  
CC TURN 88 89  
CC TURN 92 93  
CC HELIX 94 97  
CC TURN 100 101  
CC STRAND 103 106  
CC TURN 113 114  
CC HELIX 115 119

FT TURN 124 126  
FT STRAND 129 142  
FT HELIX 131 142  
FT TURN 143 144  
FT STRAND 148 167  
FT STRAND 168 168  
FT HELIX 169 180  
FT TURN 181 182  
FT STRAND 185 185  
FT HELIX 189 196  
FT TURN 197 197  
FT HELIX 198 207  
FT STRAND 212 220  
FT STRAND 222 223  
FT TURN 226 231  
FT STRAND 232 233  
FT TURN 235 239  
FT STRAND 240 240  
FT STRAND 243 243  
FT STRAND 248 248  
FT STRAND 268 268  
FT STRAND 276 277  
FT STRAND 279 280  
FT STRAND 282 284  
FT STRAND 286 289  
FT HELIX 290 292  
FT TURN 296 297  
FT STRAND 299 303  
FT HELIX 311 313  
FT STRAND 315 317  
FT TURN 319 320  
FT STRAND 322 325  
FT TURN 326 327  
FT STRAND 328 332  
FT TURN 337 338  
FT STRAND 340 344  
FT TURN 346 348  
FT HELIX 351 353  
FT STRAND 355 355  
FT STRAND 357 358  
FT STRAND 360 361  
FT STRAND 364 366  
FT TURN 367 370  
FT STRAND 371 374  
FT TURN 380 381  
FT STRAND 383 383  
FT STRAND 385 387  
FT HELIX 393 395  
FT STRAND 398 399  
FT STRAND 406 408  
FT STRAND 410 411  
FT HELIX 413 415  
FT STRAND 417 421  
FT TURN 422 423  
FT STRAND 424 428  
FT TURN 432 433  
FT HELIX 435 437  
FT STRAND 439 441  
FT TURN 443 444  
FT STRAND 447 449  
FT TURN 450 451  
FT STRAND 452 459  
FT TURN 464 464  
FT STRAND 467 472  
FT TURN 474 475  
FT HELIX 478 480  
FT STRAND 483 484  
FT TURN 486 487  
FT STRAND 490 492  
FT TURN 493 496  
FT STRAND 497 501

Query Match

25.8%; Score 245; DB 1; Length 528;

Best Local Similarity 44.8%; Pred. No. 5.6e-16;  
Matches 60; Conservative 18; Mismatches 46; Indels 10; Gaps 4

```
QY      44  INTTGATVAVOSTNTNIRAVRGRLLTTGADVAREHPVPINPRGLPIINQRTIVELVSNAEL 103
         | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db       5  IKSTSGATISOSIKQTEALRELRKRG--LHIDIPVLDDPTTLDERNNRYIIVELSNSDTE 62

QY     104 SVTLADVDYNNAVVGGRAGNASAYFFH--PDNOEDAEATHTFTDVONRYTFAFGANDRL 161
         | : | : | | | | | | | | | | : | : | : | : | : | : | : | : | : |
Db      63 SIEVGIDVTYNNAVVARACTOYSTFLRDAPSSASD----YLETGT-DOHSILPEFYGYDL 116

QY     162 EQLAGNILRENIEIG 175
         | : | | : | | | |
Db     117 ERMAHOSROOIPLG 130
```

## RESULT 9

ID	REF_ID	STANDARD	PRT	290 AA.
RT	RT	RT	RT	RT
AC	P3185	09S819		
DT	01-OCT-1993	(Rel. 27, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Ribosome inactivating protein bryodin I precursor (RNA N-glycosidase (EC 3.2.2.22) (BD1).			
OS	Bryonia dioica (Red Bryony).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.			
OC	NCBI_TaxId:3652;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).			
RC	TISSUE=Leaf;			
RX	MEDLINE=97228081; PubMed=9115985;			
RA	Gawlak S.L., Neubauer M., Kiehl H.E., Chang C.Y.Y., Einspahr H.M., Siegel J.C.B.;			
RA	"Molecular, biological, and preliminary structural analysis of recombinant bryodin 1, a ribosome-inactivating protein from the plant Bryonia dioica.";			
RT	Biochemistry 36:3095-3103(1997).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	Siegal J.C.B.;			
RT	"Cloning and expression of a gene encoding bryodin 1 from Bryonia dioica.";			
RL	Patent number US541110, 30-JUL-1996.			
RN	[3]			
RP	SEQUENCE OF 24-66.			
RC	TISSUE=Seed;			
RX	MEDLINE=89326691; PubMed=2753596;			
RA	Montecucchi P.C., Lazzerini A.M., Barbieri L., Stirpe F., Sorla M., Lappi D.;			
RA	"N-terminal sequence of some ribosome-inactivating proteins.";			
RT	Int. J. Pept. Protein Res. 33:263-267(1989).			
RL	[4]			
RN	SEQUENCE OF 24-43.			
RP	TISSUE=Root;			
RC	MEDLINE=95151812; PubMed=7849072;			
RA	Stiegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B., Marguier H.;			
RA	"Characterization of ribosome-inactivating proteins isolated from Bryonia dioica and their utility as carcinoma-reactive immunoconjugates.";			
RT	Bioconj. Chem. 5:423-429(1994).			
RL	Bioconj. Chem. 5:423-429(1994).			
CC	-1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS PROTEIN SYNTHESIS IN ANIMAL CELLS.			
CC	-1- CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.			
CC	-1- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO PRODUCE A SHORTER PROTEIN.			
CC	-1- BIOCHEMISTRY: Especially useful as immunotoxin for pharmacological applications as it has low toxicity in rats and mice but is potent once inside target cells.			

```
CC CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC CC TYPE 1 RIP SUBFAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR DR EMBL; I24020; ? NOT_ANNOTATED_CDS.
DR DR PIR; S16491; S16491.
DR DR PDB; 1BRV; 04-MAR-98.
DR DR InterPro: IPRO001574; RIP.
DR DR Pfam; PF00161; RIP; 1.
DR DR PRINTS; PR00396; SHIGARICIN.
DR DR PROSITE; PS00275; SHIGA_RICIN; 1.
KM KM Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
KW KW 3D-structure; Multigene family; Glycoprotein; Signal.
FT FT SIGNAL 1 23
FT FT CHAIN 24 270 RIBOSOME-INACTIVATING PROTEIN BROIDIN I.
FT FT PROPEP 271 290 MISSING IN MATURE PROTEIN.
FT FT ACT_SITE 183 183 BY SIMILARITY.
FT FT ACT_SITE 212 212
FT FT CARBOHYD 214 214 N-LINKED (GLCNAc...?) (POTENTIAL).
FT FT CARBOHYD 250 250 E-XK: REDUCES ACTIVITY 10-FOLD.
FT FT MUTAGEN 212 212 RSSIS -> LRHXI (IN REF. 3).
FT FT CONFLICT 61 65
FT FT STRAND 25 28
FT FT TURN 30 31
FT FT HELIX 34 46
FT FT TURN 47 47
FT FT STRAND 50 54
FT FT TURN 55 56
FT FT STRAND 57 60
FT FT HELIX 66 69
FT FT STRAND 70 76
FT FT TURN 78 79
FT FT STRAND 82 88
FT FT TURN 88 92
FT FT STRAND 93 99
FT FT TURN 100 101
FT FT STRAND 102 105
FT FT HELIX 109 114
FT FT TURN 115 117
FT FT TURN 120 121
FT FT STRAND 124 127
FT FT HELIX 134 141
FT FT TURN 142 142
FT FT HELIX 145 147
FT FT STRAND 150 150
FT FT HELIX 152 163
FT FT TURN 164 165
FT FT HELIX 167 186
FT FT STRAND 187 187
FT FT HELIX 188 196
FT FT STRAND 202 202
FT FT HELIX 206 213
FT FT TURN 214 214
FT FT HELIX 215 225
FT FT TURN 226 230
FT FT STRAND 231 239
FT FT TURN 241 242
FT FT STRAND 245 250
FT FT TURN 251 252
FT FT HELIX 254 257
FT FT TURN 258 259
FT FT STRAND 260 260
FT FT STRAND 263 263
FT FT HELIX 266 268
SQ SQ SEQUENCE 290 AA; 31788 MW; E96GCD9C031A42DB CRC64;
```

Query Match	Best Local Similarity	Score	DB 1:	Length	563:
Matches	58; Conservative	30; Mismatches	65; Indels	20; Gaps	4; :
QY	9	VIMMVAATWVLCFGSTSGMSFLEDDNNIFPKQYPLINFTTAGATVQSYTNFIRAVGRILT	68		
DB	8	MLMFYIVLALCSVGIOG-----IDYPSVSPNLGAKSATYRDLISNRKTYA	55		
QY	69	TGADVRHEIPEVLPNVRGCLPIQORFLVLSLMSHAELSTVLADYNANAYVGRAGNSATFF	128		
DB	56	TGTYEVNMLPLVLRSEVQVRSFVLPVLTNNGVTPLADVNLVYVAASGNANSITFF	115		
QY	129	HPDNOEDAEAT--THLFTDVQNNRYFAFGGVNDRLEQLAGNLRNIEILGNSP	179		
DB	116	-----KDAIEVQKSNLTFVGTGRON--TLSTGTAVNDNETAANNRRSIEIGSPSL	162		
RESULT 11					
RIPB_LUFCY	ID	STANDARD:	PRT;	250 AA.	
AC	P22851:				
DT	01-AUG-1991 (Rel. 19, last sequence update)				
DT	01-AUG-1991 (Rel. 19, last sequence update)				
DT	28-FEB-2003 (Rel. 41, last annotation update)				
DE	Ribosome-inactivating protein Luffin-B (rRNA N-glycosidase) (EC 3.2.2.22).				
OS	Luffa cylindrica (Smooth loofah) (Sponge gourd).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae;				
OC	eurosid 1; Cucurbitales; Cucurbitaceae; Luffa.				
OX	NCBI_TaxID=3670;				
RN	[1]				
RP	SEQUENCE.				
RC	TISSUE=Seed;				
RX	MEDLINE=91248488; PubMed=136866;				
RA	Islam M.R., Hirayama H., Funatsu G.;				

RT "Complete amino acid sequence of luffin-b, a ribosome-inactivating protein from sponge gourd (*Luffa cylindrica*) seeds.";  
 RL Agric. Biol. Chem. 55:229-238(1991).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 DR PIR; JN0108; JN0108.  
 DR HSSP; P16094; IABC.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin.  
 KW ACT\_SITE 160 BY SIMILARITY.  
 SQ SEQUENCE 250 AA; 27293 MW; F01ABDC8A1078700 CRC64;

Query Match 23.0%; Score 218; DB 1; Length 250;  
 Best Local Similarity 35.6%; Pred. No. 9e-14;

Matches 47; Conservative 31; Mismatches 48; Indels 6; Gaps 2;

OY 44 INFTAGATVOSTNFRVAVRGRLTGADVREIPLVLRNRYGLPINORFIIVELSNMEL 103  
 DB 3 VFSLSGADSKSYKFTALRKALPSKEKVSNIPLPSASGA--SYTILMOISNDYAK 59  
 OY 104 SYTALDVTNAVYVGYRAGNSAYFFHPDNOEDAEATHLEFDVQNRFFAFGANYDRLEQ 163  
 DB 60 ATMAIDVTNAVYIMGYVNSTSYF--ANESDAKLASGYVKGSLTVIPISGNERNLON 116  
 OY 164 LAGNIRENIEG 175  
 DB 117 MAGKIREKIPLG 128

## RESULT 12

MLA\_VISAL STANDARD; PRT; 254 AA.

AC P81446;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Beta-galactosidase specific lectin I A chain (MLA) (ML-I A) (rRNA N-glycosidase) (EC 3.2.2.22).  
 OS Viscum album (European mistletoe).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Santalales; Santalaceae; Viscum.  
 OX NCBI\_TaxID=3972;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN-Subsp. album.  
 RX MEDLINE=97134581; PubMed=8980141;  
 RA Soier M.H., Stoeva S., Schwamborn C., Wilhelm S., Stiefel T., Voelter W.;  
 RT "Complete amino acid sequence of the A chain of mistletoe lectin I.";  
 RL FEBS Lett. 399:153-157(1996).  
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNTS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RRNA. THE B CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY INTO THE CELL OF THE A CHAIN. B CHAINS ARE ALSO RESPONSIBLE FOR CELL AGGUTINATION (LECTIN ACTIVITY).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: Disulfide-linked dimer of A and B chains.  
 CC -1- PHARMACEUTICAL: Due to its immunomodulative effects it is being studied in clinical trials in cancer patients as it may slow the growth of cancer cells and be an effective treatment for solid tumors.  
 CC -1- MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA AND NON-GLYCOSYLATED FORM MLA'.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

CC TYPE 2 RIP SUBFAMILY.  
 DR PIR; PD0018; PD0018.  
 DR HSSP; P11140; IABR.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; FALSE\_NEG.  
 KW Plant defense; Hydrolase; Toxin; Repeat; Glycoprotein; Lectin.  
 KW ACT\_SITE 165 BY SIMILARITY.  
 FT CARBOHYD 112 112  
 FT VARIANT 15 15  
 FT VARIANT 66 66  
 FT VARIANT 112 112  
 FT VARIANT 116 116  
 FT VARIANT 133 134  
 FT VARIANT 140 140  
 FT VARIANT 144 144  
 FT VARIANT 151 151  
 FT VARIANT 179 179  
 FT VARIANT 184 184  
 FT VARIANT 190 190  
 FT VARIANT 218 218  
 FT VARIANT 223 224  
 FT VARIANT 231 231  
 FT VARIANT 235 235  
 SQ SEQUENCE 254 AA; 28478 MW; 53BAF9BD3E0FF6E7 CRC64;

Query Match 22.8%; Score 216.5; DB 1; Length 254;  
 Best Local Similarity 43.8%; Pred. No. 1.3e-13;

Matches 57; Conservative 24; Mismatches 40; Indels 9; Gaps 7;

OY 48 TAGATVOSTNFRVAVRGRLTGADVREIPLVLRNRYGLPIN--QRFIIIVELSNMELSY 105  
 DB 9 THQTTGGEYFRFTLLRDYVSSG--FSNELPL--KSTIPVSDAQRRVLELTNOGDV 66  
 OY 106 TLALDVTNAVYVGYRAGNSAYFFHPDNOEDAEATHLEFDVQNRFFAFGANYDRLEQ 165  
 DB 67 TAAIDVTNAVYVGYAGDQSYFLR-DAPRAE--THLFTGT-TRSSLFPGSYPDLEKRA 122  
 OY 166 GNIRENIEG 175  
 DB 123 GH-RDQIPLG 131

## RESULT 13

RIP2\_MOMBA STANDARD; PRT; 286 AA.

AC P29339;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ribosome-inactivating protein momordin II precursor (rRNA N-glycosidase) (EC 3.2.2.22).  
 OS Momordica balsamina (Bitter melon) (Bitter melon).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.  
 OX NCBI\_TaxID=3672;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seed.  
 RX MEDLINE=93027170; PubMed=1408771;  
 RA Orliga M., Better M.;  
 RT "Momordin II, a ribosome inactivating protein from Momordica balsamina, is homologous to other plant proteins.";  
 RL Nucleic Acids Res. 20:4662-4662(1992).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -







```

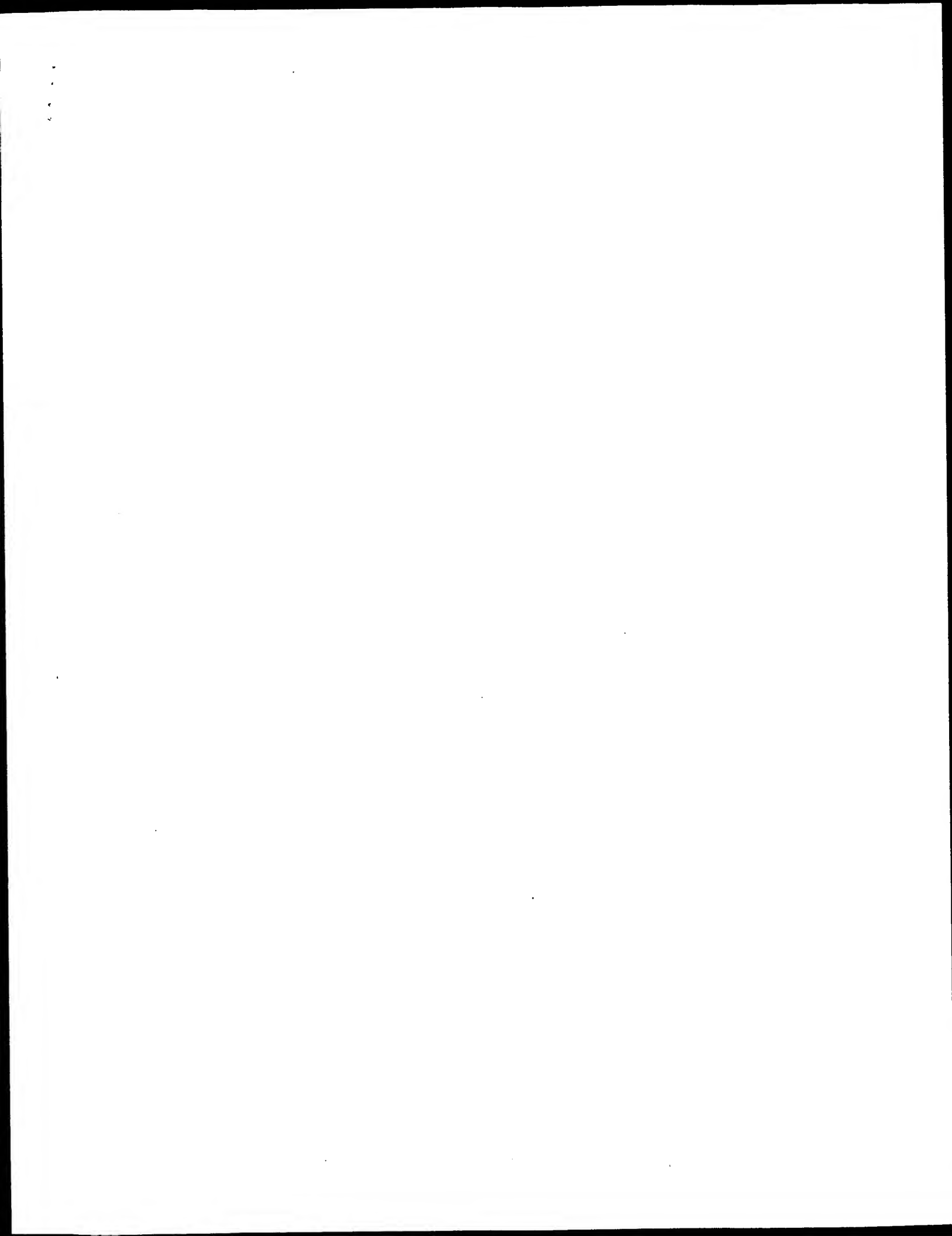
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L12243; AAA16312.1; -.
DR PIR; J0753; J0753.
DR HSSP; P09989; 1MR1.
DR InterPro; IPR001574; R1P.
DR Pfam; PF00161; R1P; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; FALSE NEG.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;
KW Glycoprotein.
FT SIGNAL 1 26
FT PROPEP 27 46 POTENTIAL.
FT CHAIN 47 297 RIBOSOME-INACTIVATING PROTEIN GELONIN.
FT PROPEP 298 316 MISSING IN MATURE PROTEIN.
FT DISULFD 90 96
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .).
FT ACT_SITE 212 212
FT CONFLICT 90 90
FT CONFLICT 93 93 C -> K (IN REF. 2).
FT CONFLICT 93 93 P -> D (IN REF. 2).
SQ SEQUENCE 316 AA; 35418 MW; 1252F3E710901B85 CRC64;

Query Match 21.9%; Score 208; DB 1; Length 316;
Best Local Similarity 36.5%; Pred. No. 1,1e-12;
Matches 69; Conservative 26; Mismatches 68; Indels 26; Gaps 12;

OY 5 GNTIVIM-YAVATWLC-----FGSTSG-WSFTLEDNNIFPKQPI--INFTAGATVQS 55
Db 3 GNMKVYWKIAVATWTFECCTIVLGSTARIPSLPTNDEEFTSKTGLDVTVSFKGATYIT 62
OY 56 YTNFIRAVRGRLTGGADVREHPIVLPNRVGLPINQRFILVELSNHAELSVTLALDVTNAY 115
Db 63 YVNFLELRVRLKPEGN-SHGIPLLRKKDDP-GKCFVLVALSNDNGQLAEIATIDVTSYV 120
OY 116 VVGYRAGNSAYFFH--PDNODAEATHTLFTD-VONRYTFAPGNYDRLEQLAGN--LRE 170
Db 121 VVGVOVNRBSYFFKADP-----AAYEGLFKNITIKTR--LHFGSIPSL---GEKAYRE 170
OY 171 NIELGNPL 179
Db 171 TTDLGIEPL 179

```

Search completed: September 16, 2003, 11:46:02  
 Job time : 7.2752 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:40:08 ; Search time 28.5521 Seconds  
(without alignments)  
1617.791 Million cell updates/sec

Title: US-10-083-336a-2

Perfect score: 948  
Sequence: 1 MKPGGNTIVIMVAVATWLC.....RLQDLAGNLRNIELNGNPL 179

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirts:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	745	78.6	541	10	Q41174
2	308.5	32.5	580	10	Q94BW3
3	305.5	32.2	581	10	Q94BW5
4	303.5	32.0	549	10	Q9FV22
5	303.5	32.0	580	10	Q94BW4
6	282	29.7	528	10	Q06076
7	265	28.0	252	10	Q38760
8	263.5	27.8	289	10	Q04KE4
9	261.5	27.6	247	10	Q94KE3
10	261.5	27.6	289	10	Q41216
11	252	26.6	563	10	Q04367
12	249.5	26.3	564	10	Q94VR2
13	249.5	26.2	270	10	Q81PV7
14	248.5	25.8	547	10	Q94KE9
15	245	25.8	251	10	Q96235
16	245	25.8	251	10	Q96237

17	245	25.8	251	10	Q96236
18	245	25.8	252	10	Q38761
19	240.5	25.4	270	10	Q41611
20	237.5	25.1	293	10	Q84542
21	236	24.9	563	10	Q94582
22	236	24.9	563	10	Q86732
23	230	24.3	278	10	Q00980
24	226.5	23.9	293	10	Q8Y100
25	226.5	23.9	604	10	Q94KE4
26	222	23.4	565	10	Q94KE5
27	221.5	23.0	299	10	Q04071
28	218.5	23.0	251	10	Q862N9
29	213	22.5	298	10	Q04358
30	211.5	22.3	264	10	Q9F5H2
31	210	22.2	566	10	Q04072
32	208.5	22.0	286	10	Q04127
33	208.5	22.0	286	10	Q9FUV7
34	206.5	21.8	249	10	Q8LK05
35	206.5	21.8	531	10	Q8RXH6
36	204	21.5	300	10	Q04356
37	203	21.4	297	10	Q8G2P0
38	202.5	21.4	254	10	Q8LK06
39	201	21.2	565	10	Q8W243
40	200.5	21.1	249	10	Q8RXH7
41	198.5	20.9	603	10	Q9M653
42	191	20.1	300	10	Q04357
43	187.5	19.8	570	10	Q22415
44	185.5	19.6	570	10	Q41358
45	185	19.5	569	10	P93543

## ALIGNMENTS

RESULT 1  
Q41174  
ID Q41174 PRELIMINARY; PRT; 541 AA.  
AC Q41174;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)  
DE Protein A chain (EC 3.2.2.22) (RMA N-glycosidase)  
DE (Fragment).  
OS Rictinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Rictinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92338377; PubMed-1633311;  
RA Roberts L.M., Tregear J.W., Lord J.M.;  
RT "Molecular cloning of rictin";  
RL Targeted Diagn. Ther. 7:81-97(1992).  
CC -I- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RNA.  
CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
CC EMBL; S40366; AAB2582.1; -.  
DR HSSP; P02879; 1886.  
DR InterPro; IPR000772; Rictin\_B\_lectin.  
DR InterPro; IPR001574; RIP.  
DR InterPro; IPR001400; Somatotropin.  
DR Pfam; PF00652; Rictin\_B\_lectin; 6.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR SMART; SM00458; RICTIN\_2.  
DR PROSITE; PS50231; RICTIN\_B\_LECTIN; 2.  
DR PROSITE; PS00275; SHIGA\_RICTIN; 1.  
DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
KW Hydrolase; Toxin.  
FT NON\_TER 1  
SQ SEQUENCE 541 AA; 60281 MW; 287B2CDEF1P2E9D9 CMC64;

Query Match 78.6%; Score 745; DB 10; Length 541;  
 Best Local Similarity 99.3%; Pred. No. 2.1e-62;  
 Matches 143; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 36 IPPKQPIINFTAGTQSYNFRVAVGRGLTGADVHEHLPVLPNRYGLPINORFIIY 95  
 |||||  
 DB 1 IPPKQPIINFTAGTQSYNFRVAVGRGLTGADVHEHLPVLPNRYGLPINORFIIY 60  
 OY 96 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNCDEDAITHLFTDVONRYTFAFG 155  
 |||||  
 DB 61 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNCDEDAITHLFTDVONRYTFAFG 120  
 OY 156 GNYDRLEQLAGNRENIEELGNGPL 179  
 |||||  
 DB 121 GNYDRLEQLAGNRENIEELGNGPL 144

## RESULT 2

O94BW3 PRELIMINARY; PRT; 580 AA.

AC O94BW3;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin III precursor  
 (EC 3.2.2.22) (RNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomin proteins and study of their expression  
 RT patterns.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC EMBL; AY039803; AAK82458.1; -  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 KW Hydrolase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 580  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN III  
 SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7FB558 CRC64;

Query Match 32.5%; Score 308.5; DB 10; Length 580;  
 Best Local Similarity 50.7%; Pred. No. 8.5e-21;  
 Matches 71; Conservative 20; Mismatches 44; Indels 5; Gaps 4;

OY 41 YPIINFTAGTQSYNFRVAVGRGLTGADVHEHLPVLPNRYGLPINORFIIYELSN 99  
 |||||  
 DB 33 YQVTFITTKNATKTSYTOFIEALRAQLASGEE-PHGIVMKERSTVPSKRFILVELSNW 91  
 OY 100 HAEISVTLADVTNAYVVGVRAGNSAYFFHPDNCDEDAITHLFTDVONRYTFAFGNYD 159  
 |||||  
 DB 92 AADSPVTLAVDTNAYVVGVRAGNSAYFFHPDNCDEDAITHLFTDVONRYTFAFGNYD 148  
 OY 160 RLEQLAGNRENIEELGNGPL 179  
 |||||  
 DB 149 DLERVAGRRERIEELGMDPL 168

RESULT 3  
 O94BW5

ID O94BW5 PRELIMINARY; PRT; 581 AA.

AC O94BW5;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin I precursor  
 (EC 3.2.2.22) (RNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomin proteins and study of their expression  
 RT patterns.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC EMBL; AY039801; AAK82458.1; -  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 5.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 KW Hydrolase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 581  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIN I  
 SQ SEQUENCE 581 AA; 64215 MW; 6EB5F5B8FBA3D196 CRC64;

Query Match 32.2%; Score 305.5; DB 10; Length 581;  
 Best Local Similarity 50.7%; Pred. No. 1.6e-20;  
 Matches 71; Conservative 19; Mismatches 45; Indels 5; Gaps 4;

OY 41 YPIINFTAGTQSYNFRVAVGRGLTGADVHEHLPVLPNRYGLPINORFIIYELSN 99  
 |||||  
 DB 33 YQVTFITTKNATKTSYTOFIEALRAQLASGEE-PHGIVMKERSTVPSKRFILVELSNW 91  
 OY 100 HAEISVTLADVTNAYVVGVRAGNSAYFFHPDNCDEDAITHLFTDVONRYTFAFGNYD 159  
 |||||  
 DB 92 AADSPVTLAVDTNAYVVGVRAGNSAYFFHPDNCDEDAITHLFTDVONRYTFAFGNYD 148  
 OY 160 RLEQLAGNRENIEELGNGPL 179  
 |||||  
 DB 149 DLERVAGRRERIEELGMDPL 168

## RESULT 4

O9FV22 PRELIMINARY; PRT; 549 AA.

AC O9FV22;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Type II ribosome-inactivating protein cinnamomin (EC 3.2.2.22) (RNA  
 DE N-glycosidase) (Fragment).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Xie L., Liu W.-Y., Wang E.-D.;  
 RT "Molecular cloning of cinnamomin A-, B-chain and the expression,  
 RT purification, characterization and mutagenesis of the A-chain.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.

41 YPINFETAGTVOSYTNFIRAVRGRI

DT	01-NOV-1998	(TREMBLrel. 08, last seq
----	-------------	--------------------------

DT	01-NOV-1998	(TREMBLrel. 08, last seq
----	-------------	--------------------------

DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE ABRIN-E (RNA N-glycosidase) (EC 3.2.2.22) (Fragment).  
 GN RIP.  
 OS Abrus precatorius (Indian liquorice) (Crab's eye).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RA MEDLINE=91201329; PubMed=2016300;  
 RX Evensen G., Mathiesen A., Sundan A.;  
 RT "Direct molecular cloning and expression of two distinct abrin A-  
 chains".  
 RL J. Biol. Chem. 266:6848-6852(1991).  
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
 CC ABRIN-A IS MORE TOXIC THAN RICIN.  
 CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
 FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
 PRECEDES ENDOCYTOSIS.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA)  
 CC -1- SIMILARITY: TO OTHER BACTERIAL AND PLANT RIBOSOME-INACTIVATING  
 PROTEINS. BELONGS TO TYPE 2 RIP.  
 CC EMBL; X54872; CAA38654.1; -  
 DR HSSP; P11140; IABR.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_TOXIN; Repeat; Glycoprotein; Lectin.  
 KW Hydroxylase; Glycosidase; Toxin; Repeat; Glycoprotein; Lectin.  
 FT CHAIN 1 252 ABRIN E, A CHAIN (BY SIMILARITY).  
 FT NON\_TER 252 252  
 SQ SEQUENCE 252 AA; 28309 MW; BRF0846B9E92B5DE CRC64;  
  
 Query Match 28.0%; Score 265; DB 10; Length 252;  
 Best Local Similarity 47.7%; Pred. No. 3.9e-17;  
 Matches 63; Conservative 16; Mismatches 47; Indels 6; Gaps 3;  
  
 QY 44 INFTAGATVOSTYNTNFRVAVRGRLTGADVHEIPLVLPNRYGLPIINORFIVELSNHAEI 103  
 DB 6 IKFTSEGATSSYKOFIALERLRG--LIHDIPLVLRDPTVEERNRYIVELSNSERE 63  
 QY 104 SVTALDVTNAVYVVGIRAGNSAYFFHPDNOEDAEIHLFTDVONRTTFAGGNYDLQ 163  
 DB 64 SIEVGIDVTNAVYVVRAGSOSYFL--DDAPASASTYLFVGTQ-RYSLRFDGSGDLER 119  
 QY 164 LAGNLRNIEIG 175  
 DB 120 WAHOTREQISUG 131  
  
 RESULT 8  
 Q94KE4 PRELIMINARY; PRT; 289 AA.  
 AC Q94KE4;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)  
 DE Trichosanthin precursor (EC 3.2.2.22) (RNA N-glycosidase).  
 GN TCS.  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 NCBI\_TaxID=3677;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Yuan H., Wang Y., An C., Chen Z.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC EMBL; AF367252; AAK52960.1; -  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_TOXIN; 1.  
 KW Hydroxylase; Signal; Toxin.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 270 TRICHOSANTHIN.  
 SQ SEQUENCE 289 AA; 31706 MW; A6D5602549CA5657 CRC64;  
  
 Query Match 27.8%; Score 263.5; DB 10; Length 289;  
 Best Local Similarity 42.1%; Pred. No. 6.5e-17;  
 Matches 56; Conservative 32; Mismatches 38; Indels 7; Gaps 4;  
  
 QY 44 INFTAGATVOSTYNTNFRVAVRGRLTGADVHEIPLVLPNRYGLPIINORFIVELSNHAEI 103  
 DB 25 VEFRLSGATSSSYGVFISLNRKALPNERKL-YDIPLT--RSSLPQSQRALIHITNVADE 81  
 QY 104 SVTALDVTNAVYVVGIRAGNSAYFFHPDNOEDAEIHLFTDVONRTTFAGGNYDLQ 162  
 DB 82 TTSVADVTNAVYVIMYGRAGDTSYFF--NEASATFAAKYFRKDSMRKITLIPYSGNYERLO 138  
 QY 163 OLGNLRNIEIG 175  
 DB 139 TAAKIRENIPIG 151  
  
 RESULT 9  
 Q9LRE3 PRELIMINARY; PRT; 247 AA.  
 AC Q9LRE3;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, last annotation update)  
 DE Trichobakin (EC 3.2.2.22) (RNA N-glycosidase) (Fragment).  
 GN TPK.  
 OS Trichosanthes sp. Bac Kan 8-98.  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 NCBI\_TaxID=118182;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Phan Van Chi, Nguyen Thuy Ha, Le Tran Binh;  
 RT "genomic DNA clone for mature typ-1 ribosome-inactivating protein from  
 Trichosanthes sp. sample 01 Bac Kan 8-98 (Hanoi)".  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC EMBL; AB039324; BAA92530.1; -  
 DR HSSP; P09989; IMRJ.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_TOXIN; 1.  
 KW Hydroxylase; Toxin.  
 FT NON\_TER 1 1  
 FT NON\_TER 247 247  
 SQ SEQUENCE 247 AA; 27199 MW; 89811AC32892F03F CRC64;  
  
 Query Match 27.6%; Score 261.5; DB 10; Length 247;  
 Best Local Similarity 42.9%; Pred. No. 8.1e-17;  
 Matches 57; Conservative 32; Mismatches 37; Indels 7; Gaps 4;  
  
 QY 44 INFTAGATVOSTYNTNFRVAVRGRLTGADVHEIPLVLPNRYGLPIINORFIVELSNHAEI 103  
 DB 11 INFTAGATVOSTYNTNFRVAVRGRLTGADVHEIPLVLPNRYGLPIINORFIVELSNHAEI 103

Db 2 VSFRLSGATSSSYGVFISNLRKALPYERKL-YDIPLL--RSTLPSCQRALHILTNVADE 58  
 QY 104 SVTLADVTNAYVYGVFAGNSAVFFHPDNOEDA-FAITHLFTDVQNRRTFFGNGYDRLE 162  
 DB 59 TISVAIDVTNAYVYGVFAGNSAVFFHPDNOEDA-FAITHLFTDVQNRRTFFGNGYDRLE 115  
 QY 163 QLAGNLRNIEIG 175  
 DB 116 IAAGKIRENIPLG 128

## RESULT 10

041216 PRELIMINARY: PRT: 289 AA.  
 ID 041216

AC 041216;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
 DE Trichosanthin (EC 3.2.2.22) (rRNA N-glycosidase).  
 GN TRICHOSANTHIN, TCS.  
 OS Trichosanthin killowii (Mongolian snake-gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 OX NCBI\_TaxID=3677;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-94271613; PubMed=8003348;  
 RX Zheng H., Wang B., Shaw P., Yeung H.;  
 RT [Cloning and DNA sequencing of the gene encoding trichosanthin].";  
 RT I Chuan Hsueh Pao 21:42-51(1984).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: S70176; AAB31048.1; -;  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP.1  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KM Hydrolase; toxin.  
 SQ SEQUENCE 289 AA; 31650 MW; 286AC14D48BCA175 CRC64;

Query Match 27.6%; Score 261.5; DB 10; Length 289;  
 Best Local Similarity 42.1%; Pred. No. 1e-16;  
 Matches 56; Conservative 32; Mismatches 38; Indels 7; Gaps 4;

QY 44 INFTAGATVQSYTNFIRAVRGRLTGADVREHPVLPNVRGPIPNQRFILVELSNHAEI 103  
 DB 25 VSFRLSGATSSSYGVFISNLRKALPYERKL-YDIPLL--RSTLPSCQRALHILTNVADE 81  
 QY 104 SVTLADVTNAYVYGVFAGNSAVFFHPDNOEDA-FAITHLFTDVQNRRTFFGNGYDRLE 162  
 DB 82 TISVAIDVTNAYVYGVFAGNSAVFFHPDNOEDA-FAITHLFTDVQNRRTFFGNGYDRLE 115  
 QY 163 QLAGNLRNIEIG 175  
 DB 139 IAAGKIRENIPLG 128

## RESULT 11

004367 PRELIMINARY: PRT: 563 AA.  
 ID 004367

AC 004367;  
 DT 01-JUL-1997 (TREMblrel. 04, Created)  
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)  
 DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
 DE Ribosome inactivating protein precursor (EC 3.2.2.22) (rRNA  
 DE N-glycosidase).  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.

OX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-98112023; PubMed=9450339;  
 RA Van Damme E.J., Roy S., Barre A., Rouge P., Van Leuven F.,  
 RA Peumans W.J.;  
 RT "The major elderberry (Sambucus nigra) fruit protein is a lectin  
 RT derived from a truncated type 2 ribosome-inactivating protein.";  
 RL Plant J. 12:1251-1260(1997).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: U76524; AAC15886.1; -;  
 DR HSSP: P02879; ZAAI.  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP.1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Signal; Toxin.  
 FT SIGNAL 1 25  
 FT CHAIN 26 297 POTENTIAL.  
 FT CHAIN 298 563 RIBOSOME INACTIVATING PROTEIN, A CHAIN.  
 SQ SEQUENCE 563 AA; 62336 MW; 3ED2B6C08E796205 CRC64;

Query Match 26.7%; Score 253; DB 10; Length 563;  
 Best Local Similarity 35.3%; Pred. No. 1.6e-15;  
 Matches 61; Conservative 33; Mismatches 59; Indels 20; Gaps 4;

QY 9 VIMAVAVATWLCFSTSGMSTFLDNNIFPKQYPIINFTAGATVQSYTNFIRAVRGRL 68  
 DB 8 MLYTIVVLAICSVGIQ-----IDPVSFNLGAKSAFYRDLKLNLTIVA 55  
 QY 69 TGADVREHPVLPNVRGPIPNQRFILVELSNHAEISVTALADVTNAYVYGVFAGNSAVFF 128  
 DB 56 TGYEVNCLPVLRESEVQKRNREVLVLTNMGNTVTLAVDTNLYVVAFSNANASYFF 115  
 QY 129 HPDNOEDA-FAITHLFTDVQNRRTFFGNGYDRLEQLAGNLRNIEIGNGPL 179  
 DB 116 -----KDATOLQKSNLFVGR-QHTLPFTGNYNDELTAAGTRESIELDSPSL 162

## RESULT 12

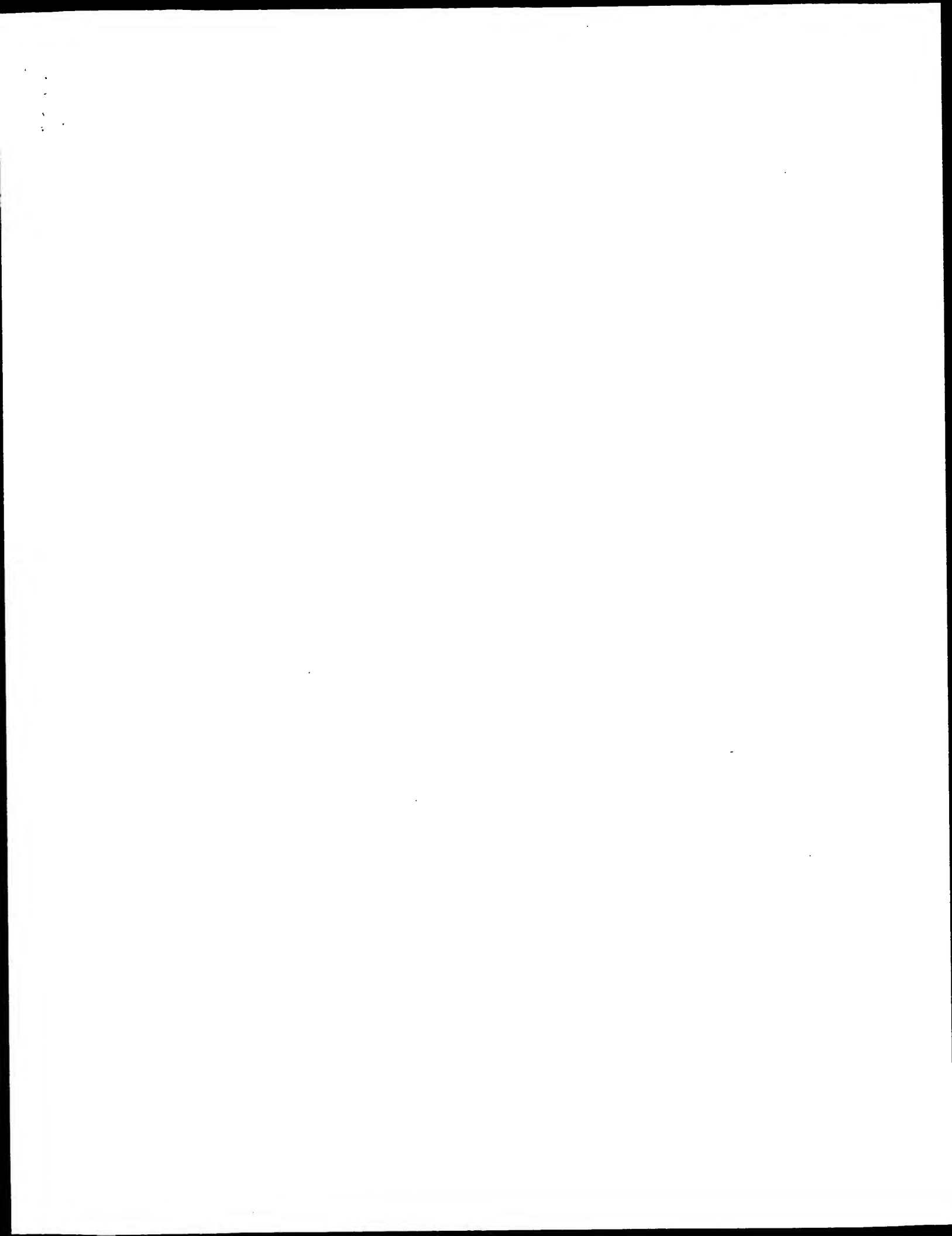
09AVR2 PRELIMINARY: PRT: 564 AA.  
 ID 09AVR2

AC 09AVR2;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DE 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Ribosome-inactivating protein precursor (EC 3.2.2.22) (rRNA  
 DE N-glycosidase).  
 OS Sambucus ebulus.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=28503;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=Leaf;  
 RA Girbes T., Iglesias R., Perez Y., Ferreras J.M., Citores L.;  
 RT "Molecular cloning of ebulin 1.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: A400822; CAC33178.1; -;  
 DR HSSP: P02879; ZAAI.  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.

[illegible]



OC	Euryzoa: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eurosid 1; Finales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX	NCBI_TaxId=3816;
RM	[1]
RP	SEQUENCE FROM N.A.
RP	TISSUE=Seed;
RC	MEDLINE=94139756; PubMed=8307038;
RX	Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT	"Cloning and expression of three abrin A-chains and their mutants
RT	derived by site-specific mutagenesis in <i>Escherichia coli</i> .";
RL	Eur. J. Biochem. 219:83-87(1994).
CC	-1- CATALYTIC ACTIVITY: ENDOPOLYESTERASE OF THE N-GLYCOSIDIC BOND AT ONE
CC	SPECIFIC ACIDONOUS ON THE 28S RRNA.
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR	EMBL: X76644; CAA54092.1; ..
DR	HSSP: P11140; IABR.
DR	Interpro: IPR001574; RIP.
DR	Pfam: PF00161; RIP; 1.
DR	PRINTS: PR00396; SHIGARICIN.
KM	Hydrolase; Toxin.
FT	NON_TER
FT	NON_TER
SO	SEQUENCE
	251 AA; 27997 MW; 3B60351839AEFB7E CRC64;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:42:03 ; Search time 11.2796 Seconds  
(without alignments)  
742.718 Million cell updates/sec

Title: US-10-083-336A-3

Perfect score: 1020

Sequence: 1 IFPKQYPIINFTAGATVQS.....ARFQYEGEMRTRIRYRNR 198

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCrUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1020	100.0	267	1	US-07-901-707-1
2	1020	100.0	267	1	US-07-988-430-1
3	1020	100.0	267	1	US-08-425-336-1
4	1020	100.0	267	1	US-08-468-1138-1
5	1020	100.0	267	1	US-08-477-484B-1
6	1020	100.0	267	1	US-08-477-484B-1
7	1020	100.0	267	3	US-08-646-360-1
8	1020	100.0	267	3	US-08-839-765-1
9	1020	100.0	267	4	US-09-136-389-1
10	1020	100.0	267	5	US-09-610-838-1
11	1020	100.0	267	5	PCT-US92-09487-1
12	1020	100.0	267	2	US-08-356-786-8
13	1020	100.0	290	1	US-08-378-761A-27
14	1020	100.0	290	1	US-08-485-286-27
15	1020	100.0	534	2	US-08-356-786-10
16	1010	99.0	267	2	US-08-218-303-16
17	1010	99.0	267	2	US-08-338-7930-61
18	1010	99.0	267	4	US-09-538-873-1
19	930.5	91.2	540	1	US-08-378-761A-77
20	930.5	91.2	540	1	US-08-485-286-77
21	342	33.5	247	1	US-08-488-1138-6
22	342	33.5	247	2	US-08-477-484B-6
23	342	33.5	247	2	US-08-466-360-6
24	342	33.5	247	3	US-08-839-765-6
25	342	33.5	247	3	US-09-136-389-6
26	342	33.5	247	4	US-09-610-838-6
27	342	33.5	267	1	US-08-378-761A-74

28	342	33.5	267	1	US-08-485-286-74	Sequence 74, Appl
29	342	33.5	289	1	US-07-923-692C-4	Sequence 4, Appl
30	342	33.5	289	1	US-08-184-237-4	Sequence 4, Appl
31	342	33.5	289	2	US-08-482-920-4	Sequence 4, Appl
32	342	33.5	289	3	US-08-484-341-4	Sequence 4, Appl
33	342	33.5	289	3	US-08-483-502-4	Sequence 4, Appl
34	342	33.5	289	4	US-09-726-551A-4	Sequence 4, Appl
35	341.5	33.5	282	1	US-08-324-301-15	Sequence 15, Appl
36	329.5	32.3	250	1	US-08-378-761A-71	Sequence 71, Appl
37	329.5	32.3	251	4	US-08-485-286-71	Sequence 71, Appl
38	323.5	31.7	251	4	US-09-538-873-3	Sequence 3, Appl
39	312.5	30.6	255	1	US-07-901-707-6	Sequence 6, Appl
40	312.5	30.6	255	1	US-07-988-430-6	Sequence 6, Appl
41	312.5	30.6	255	1	US-08-425-336-6	Sequence 6, Appl
42	312.5	30.6	255	5	PCT-US92-09487-6	Sequence 6, Appl
43	312	30.6	248	3	US-08-902-486-7	Sequence 7, Appl
44	312	30.6	290	1	US-08-245-754A-2	Sequence 2, Appl
45	312	30.6	290	2	US-08-597-731-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1

US-07-901-707-1

Sequence 1, Application US/07901707

Patent No. 5376546

GENERAL INFORMATION:

APPLICANT: Bernhardt, Susan L.

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Steve F.

APPLICANT: Lane, Julie A.

TITLE OF INVENTION: Materials Comprising and Methods of

NUMBER OF SEQUENCES: 57

TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

ADDRESSEE: Bicknell

STREET: Two First National Plaza, 20 South Clark

STREET: Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/901,707

FILING DATE: 19920619

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: No. 5376546and, Greta E.

REGISTRATION NUMBER: 35,502

REFERENCE/DOCKET NUMBER: 27129/30910

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 346-5750

TELEFAX: (312) 984-5750

TELEX: 25-3856

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-901-707-1

Query Match

100.0%; Score 1020; DB 1; Length 267;

Best Local Similarity 100.0%; Pred. No. 1.7e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVREHIEPVLPRVGLPINQRFILV 60  
Db 1 IFFKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVREHIEPVLPRVGLPINQRFILV 60

QY 61 ELSNHAELSVTLALDVNTAVVGYRAGNSAYFFHPDNOEDAFAITHLFTDVQNRRTFAFG 120  
Db 61 ELSNHAELSVTLALDVNTAVVGYRAGNSAYFFHPDNOEDAFAITHLFTDVQNRRTFAFG 120

QY 121 GNYDRLEQLAGNLRNIEELGNGPLBEAISALYYSTGTOPLTARSFIIICIMISEAR 180  
Db 121 GNYDRLEQLAGNLRNIEELGNGPLBEAISALYYSTGTOPLTARSFIIICIMISEAR 180

QY 181 FOYIEGEMRTIRIRYNRS 198  
Db 181 FOYIEGEMRTIRIRYNRS 198

## RESULT 2

US-07-988-430-1  
; Sequence 1, Application US/07988430  
; Patent No. 5416202

## GENERAL INFORMATION:

APPLICANT: Bernhard, Susan L.  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Lane, Julie A.  
APPLICANT: Lei, Shau-Ping  
TITLE OF INVENTION: Materials Comprising and Methods of  
TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/988,430  
FILING DATE: 19921209  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5416202and, Greta E.  
REGISTRATION NUMBER: 35302  
REFERENCE/DOCKET NUMBER: 31133  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-988-430-1

Query Match 100.0%; Score 1020; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVREHIEPVLPRVGLPINQRFILV 60  
Db 1 IFFKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVREHIEPVLPRVGLPINQRFILV 60

QY 61 ELSNHAELSVTLALDVNTAVVGYRAGNSAYFFHPDNOEDAFAITHLFTDVQNRRTFAFG 120  
Db 61 ELSNHAELSVTLALDVNTAVVGYRAGNSAYFFHPDNOEDAFAITHLFTDVQNRRTFAFG 120

QY 121 GNYDRLEQLAGNLRNIEELGNGPLBEAISALYYSTGTOPLTARSFIIICIMISEAR 180  
Db 121 GNYDRLEQLAGNLRNIEELGNGPLBEAISALYYSTGTOPLTARSFIIICIMISEAR 180

QY 181 FOYIEGEMRTIRIRYNRS 198  
Db 181 FOYIEGEMRTIRIRYNRS 198

## RESULT 3

US-08-425-336-1  
; Sequence 1, Application US/08425336  
; Patent No. 5621083

## GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
APPLICANT: Studinka, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/425,336  
FILING DATE: 18-APR-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Thomas C.  
REGISTRATION NUMBER: P-36,989  
REFERENCE/DOCKET NUMBER: 31394  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-425-336-1

Query Match 100.0%; Score 1020; DB 1; Length 267;

Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 IFFKQYPIINFTAGATVOSTYNTFIRAVRGRLTTGADVREHIEPVLPRNGVLPINORITLV 60
   |||||||
Db 1 IFFKQYPIINFTAGATVOSTYNTFIRAVRGRLTTGADVREHIEPVLPRNGVLPINORITLV 60
   |||||||
QY 61 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAG 120
   |||||||
Db 61 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAG 120
   |||||||
QY 121 GNYDRLEQLAGNLRNTEIELGNGPLEEASALYYSTGTGTPPLARSFTICQMISEAR 180
   |||||||
Db 121 GNYDRLEQLAGNLRNTEIELGNGPLEEASALYYSTGTGTPPLARSFTICQMISEAR 180
   |||||||
QY 181 FOYIEGEMRTIRIRNRS 198
   |||||||
Db 181 FOYIEGEMRTIRIRNRS 198

```

## RESULT 4

US-08-488-113B-1  
; Sequence 1, Application US/08488113B  
; Patent No. 5744580

## GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,113B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

## ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248

INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-488-113B-1

Query Match 100.0%; Score 1020; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 IFFKQYPIINFTAGATVOSTYNTFIRAVRGRLTTGADVREHIEPVLPRNGVLPINORITLV 60
   |||||||
Db 1 IFFKQYPIINFTAGATVOSTYNTFIRAVRGRLTTGADVREHIEPVLPRNGVLPINORITLV 60
   |||||||
QY 61 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAG 120
   |||||||
Db 61 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAG 120
   |||||||
QY 121 GNYDRLEQLAGNLRNTEIELGNGPLEEASALYYSTGTGTPPLARSFTICQMISEAR 180
   |||||||
Db 121 GNYDRLEQLAGNLRNTEIELGNGPLEEASALYYSTGTGTPPLARSFTICQMISEAR 180
   |||||||
QY 181 FOYIEGEMRTIRIRNRS 198
   |||||||
Db 181 FOYIEGEMRTIRIRNRS 198

```

## RESULT 5

US-08-477-484B-1  
; Sequence 1, Application US/08477484B  
; Patent No. 5756699

## GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,484B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

## ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248

TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-477-484B-1

Query Match 100.0%; Score 1020; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFKQYPIINFTAGATVQSTYTNFIRAVRGRLLTGADVRHEIPVLPNVRGLPINQRFILV 60  
DB 1 IFFKQYPIINFTAGATVQSTYTNFIRAVRGRLLTGADVRHEIPVLPNVRGLPINQRFILV 60  
QY 61 ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATLHFTDVQNRRTFAFG 120  
DB 61 ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATLHFTDVQNRRTFAFG 120  
QY 121 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYSTGSGTQPLTARSFIICMISEAR 180  
DB 121 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYSTGSGTQPLTARSFIICMISEAR 180  
QY 181 FOYIEGEMRTIRIRYNRRS 198  
DB 181 FOYIEGEMRTIRIRYNRRS 198

## RESULT 6

US-08-646-360-1  
Sequence 1, Application US/08646360  
Patent No. 5837491

## GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996

## CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993

## COMPUTER READABLE FORM:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Monicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-360-1

Query Match 100.0%; Score 1020; DB 2; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFKQYPIINFTAGATVQSTYTNFIRAVRGRLLTGADVRHEIPVLPNVRGLPINQRFILV 60  
DB 1 IFFKQYPIINFTAGATVQSTYTNFIRAVRGRLLTGADVRHEIPVLPNVRGLPINQRFILV 60  
QY 61 ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATLHFTDVQNRRTFAFG 120  
DB 61 ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATLHFTDVQNRRTFAFG 120  
QY 121 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYSTGSGTQPLTARSFIICMISEAR 180  
DB 121 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYSTGSGTQPLTARSFIICMISEAR 180  
QY 181 FOYIEGEMRTIRIRYNRRS 198  
DB 181 FOYIEGEMRTIRIRYNRRS 198

## RESULT 7

US-08-839-765-1  
Sequence 1, Application US/08839765  
Patent No. 6146631

## GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,765  
FILING DATE: 15-APR-1997

## CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430

;; FILING DATE: 09-DEC-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/901,707  
;; FILING DATE: 19-JUN-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/787,567  
;; FILING DATE: 04-NOV-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: McNicholas, Janet M.  
;; REGISTRATION NUMBER: 32,918  
;; REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312/707-8889  
;; TELEFAX: 312/707-9155  
;; TELEX: 650 388-1248  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 267 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-839-765-1

Query Match 100.0%; Score 1020; DB 3; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEPKQYPIINFTAGATVQSYTNFIRAVRGRLTGADVHHEIPVLPNRYGLPINOREFLV 60  
DB 1 IEPKQYPIINFTAGATVQSYTNFIRAVRGRLTGADVHHEIPVLPNRYGLPINOREFLV 60  
QY 61 ELSNHAELSTYTLADVTNAYVVGGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAFG 120  
DB 61 ELSNHAELSTYTLADVTNAYVVGGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAFG 120  
QY 121 GNYDRLEQLAGNIRENIEELGNGPLEEPAISALYYSTGQTOLPTLARSFTICOMISEAR 180  
DB 121 GNYDRLEQLAGNIRENIEELGNGPLEEPAISALYYSTGQTOLPTLARSFTICOMISEAR 180  
QY 181 FOYIEGEMRTIRIRYNRS 198  
DB 181 FOYIEGEMRTIRIRYNRS 198

RESULT 8  
US-09-136-389-1  
;; Sequence 1, Application US/09136389  
;; Patent No. 6146850  
;; GENERAL INFORMATION:  
;; APPLICANT: Better, Marc D.  
;; APPLICANT: Carroll, Stephen F.  
;; APPLICANT: Studnika, Gary M.  
;; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
;; TITLE OF INVENTION: Proteins  
;; NUMBER OF SEQUENCES: 173  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
;; STREET: 500 West Madison Street, 34th floor  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60661  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; OPERATING SYSTEM: IBM PC compatible  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/136,389  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/646,360

;; FILING DATE: 13-MAY-1996  
;; APPLICATION NUMBER: PCT/US94/05348  
;; FILING DATE: 12-MAY-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/064,691  
;; FILING DATE: 12-MAY-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/988,430  
;; FILING DATE: 09-DEC-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/901,707  
;; FILING DATE: 19-JUN-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/787,567  
;; FILING DATE: 04-NOV-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: McNicholas, Janet M.  
;; REGISTRATION NUMBER: 32,918  
;; REFERENCE/DOCKET NUMBER: 200-70.P4  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312/707-8889  
;; TELEFAX: 312/707-9155  
;; TELEX: 650 388-1248  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 267 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-09-136-389-1

Query Match 100.0%; Score 1020; DB 3; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.7e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEPKQYPIINFTAGATVQSYTNFIRAVRGRLTGADVHHEIPVLPNRYGLPINOREFLV 60  
DB 1 IEPKQYPIINFTAGATVQSYTNFIRAVRGRLTGADVHHEIPVLPNRYGLPINOREFLV 60  
QY 61 ELSNHAELSTYTLADVTNAYVVGGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAFG 120  
DB 61 ELSNHAELSTYTLADVTNAYVVGGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAFG 120  
QY 121 GNYDRLEQLAGNIRENIEELGNGPLEEPAISALYYSTGQTOLPTLARSFTICOMISEAR 180  
DB 121 GNYDRLEQLAGNIRENIEELGNGPLEEPAISALYYSTGQTOLPTLARSFTICOMISEAR 180  
QY 181 FOYIEGEMRTIRIRYNRS 198  
DB 181 FOYIEGEMRTIRIRYNRS 198

RESULT 9  
US-09-610-838-1  
;; Sequence 1, Application US/09610838  
;; Patent No. 6376217  
;; GENERAL INFORMATION:  
;; APPLICANT: Better, Marc D.  
;; APPLICANT: Carroll, Stephen F.  
;; APPLICANT: Studnika, Gary M.  
;; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
;; TITLE OF INVENTION: Proteins  
;; NUMBER OF SEQUENCES: 173  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
;; STREET: 500 West Madison Street, 34th floor  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60661  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/610,838
FILING DATE: 06-JUL-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/136,389
FILING DATE: 18-AUG-1998
APPLICATION NUMBER: 08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70. P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-610-838-1

```

```

Query Match      100.0%; Score 1020; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.7e-111;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINORFTLV 60
DB 1 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINORFTLV 60
QY 61 ELSNHAELSVTLALDVTNAYVVGGRAGNSAVFFHPDNOEDAEATHLFTDVONRRTFAFG 120
DB 61 ELSNHAELSVTLALDVTNAYVVGGRAGNSAVFFHPDNOEDAEATHLFTDVONRRTFAFG 120
QY 121 GNYDRLEQLAGNLRNIEELGNGPLLEAISALYYSTGCTOLPTLARSFIIICQMISEAR 180
DB 121 GNYDRLEQLAGNLRNIEELGNGPLLEAISALYYSTGCTOLPTLARSFIIICQMISEAR 180
QY 181 FOYIEGEMRTIRIRNRS 198
DB 181 FOYIEGEMRTIRIRNRS 198

```

```

RESULT 10
PCT-US92-09487-1

```

```

Sequence 1, Application PC/TUS9209487
GENERAL INFORMATION:
APPLICANT: Bernhard, Susan L.
APPLICANT: Better, Marc D.
APPLICANT: Caroll, Stephen F.
APPLICANT: Lane, Julie A.
APPLICANT: Lei, Shau-Ping
TITLE OF INVENTION: Materials Comprising and Methods of
Preparation and Use for Ribosome-Inactivating Proteins

```

```

NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09487
FILING DATE: 19921104
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Greta E.
REGISTRATION NUMBER: 35302
REFERENCE/DOCKET NUMBER: 31133
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-09487-1

```

```

Query Match      100.0%; Score 1020; DB 5; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.7e-111;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINORFTLV 60
DB 1 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINORFTLV 60
QY 61 ELSNHAELSVTLALDVTNAYVVGGRAGNSAVFFHPDNOEDAEATHLFTDVONRRTFAFG 120
DB 61 ELSNHAELSVTLALDVTNAYVVGGRAGNSAVFFHPDNOEDAEATHLFTDVONRRTFAFG 120
QY 121 GNYDRLEQLAGNLRNIEELGNGPLLEAISALYYSTGCTOLPTLARSFIIICQMISEAR 180
DB 121 GNYDRLEQLAGNLRNIEELGNGPLLEAISALYYSTGCTOLPTLARSFIIICQMISEAR 180
QY 181 FOYIEGEMRTIRIRNRS 198
DB 181 FOYIEGEMRTIRIRNRS 198

```

```

RESULT 11
US-08-356-786-8

```

```

Sequence 8, Application US/08356786
Patent No. 5877305
GENERAL INFORMATION:
APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
Marker

```



```

; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Pesta, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-786-8

Query Match      100.0%; Score 1020; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.7e-111;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 IPPKOYPIINFTAGATVOSTYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINORITLV 60
DB      2 IPPKOYPIINFTAGATVOSTYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINORITLV 61
QY      61 ELSNHAELSVTLADLVNTAAYVGYRAGNSATFFHPDNOEDAETHLFTVONRYTFAPG 120
DB      62 ELSNHAELSVTLADLVNTAAYVGYRAGNSATFFHPDNOEDAETHLFTVONRYTFAPG 121
QY      121 GNYDLLEQLAGNLRENIELGNGPLEEAIISALYYSTGTGTOPLTARSFFICMISPAAR 180
DB      122 GNYDLLEQLAGNLRENIELGNGPLEEAIISALYYSTGTGTOPLTARSFFICMISPAAR 181
QY      181 FOYIEGEMRTIRRYNRRS 198
DB      182 FOYIEGEMRTIRRYNRRS 199

RESULT 12
US-08-378-761A-27
; Sequence 27, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN

```

```

; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 290 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-761A-27

```

```

Query Match      100.0%; Score 1020; DB 1; Length 290;
Best Local Similarity 100.0%; Pred. No. 1.9e-111;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 IPPKOYPIINFTAGATVOSTYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINORITLV 60
DB      25 IPPKOYPIINFTAGATVOSTYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINORITLV 84
QY      61 ELSNHAELSVTLADLVNTAAYVGYRAGNSATFFHPDNOEDAETHLFTVONRYTFAPG 120
DB      85 ELSNHAELSVTLADLVNTAAYVGYRAGNSATFFHPDNOEDAETHLFTVONRYTFAPG 144
QY      121 GNYDLLEQLAGNLRENIELGNGPLEEAIISALYYSTGTGTOPLTARSFFICMISPAAR 180
DB      145 GNYDLLEQLAGNLRENIELGNGPLEEAIISALYYSTGTGTOPLTARSFFICMISPAAR 204
QY      181 FOYIEGEMRTIRRYNRRS 198
DB      205 FOYIEGEMRTIRRYNRRS 222

RESULT 13
US-08-485-286-27
; Sequence 27, Application US/08485286
; Patent No. 5646026
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/485, 286  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/378761  
FILING DATE: 26-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 290 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-286-27

Query Match 100.0%; Score 1020; DB 1; Length 290;  
Best Local Similarity 100.0%; Pred. No. 1.9e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTTGADVREHIEPVLNRYGLPINORFIIV 60  
DB 25 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTTGADVREHIEPVLNRYGLPINORFIIV 84  
OY 61 ELSNHAELSVTLADVTNAVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAFG 120  
DB 85 ELSNHAELSVTLADVTNAVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAFG 144  
OY 121 GNYDRLEQLAGNLENIELGNPLEEASISALYYSTGTQDPLTARSFIICMISEAR 180  
DB 145 GNYDRLEQLAGNLENIELGNPLEEASISALYYSTGTQDPLTARSFIICMISEAR 204  
OY 181 FOYTEGEMRTIRIRNRRS 198  
DB 205 FOYTEGEMRTIRIRNRRS 222

RESULT 14  
5248606-4  
Patent No. 5248606  
APPLICANT: WALSH, TERENCE A.; HEY, TIMOTHY D.; MORGAN,  
ALICE E.R.  
TITLE OF INVENTION: DNA ENCODING INACTIVE PRECURSOR AND  
ACTIVE FORMS OF MAIZE RIBOSOME INACTIVATIN  
NUMBER OF SEQUENCES: 49  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/535, 636  
FILING DATE: 11-JUN-1990  
SEQ ID NO: 4  
LENGTH: 290  
5248606-4

Query Match 100.0%; Score 1020; DB 6; Length 290;  
Best Local Similarity 100.0%; Pred. No. 1.9e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTTGADVREHIEPVLNRYGLPINORFIIV 60  
DB 25 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTTGADVREHIEPVLNRYGLPINORFIIV 84  
OY 61 ELSNHAELSVTLADVTNAVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAFG 120  
DB 85 ELSNHAELSVTLADVTNAVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAFG 144  
OY 121 GNYDRLEQLAGNLENIELGNPLEEASISALYYSTGTQDPLTARSFIICMISEAR 180  
DB 145 GNYDRLEQLAGNLENIELGNPLEEASISALYYSTGTQDPLTARSFIICMISEAR 204

OY 181 FOYTEGEMRTIRIRNRRS 198  
DB 205 FOYTEGEMRTIRIRNRRS 222

RESULT 15  
US-08-356-786-10  
Sequence 10, Application US/08356786  
Patent No. 5877305  
GENERAL INFORMATION:  
APPLICANT: Huston, James S.  
APPLICANT: Oppermann, Hermann  
APPLICANT: Houston, L. L.  
APPLICANT: Ring, David B.  
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
TITLE OF INVENTION: Marker  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356, 786  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/831,967  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-053  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 534 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-356-786-10

Query Match 100.0%; Score 1020; DB 2; Length 534;  
Best Local Similarity 100.0%; Pred. No. 4.8e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

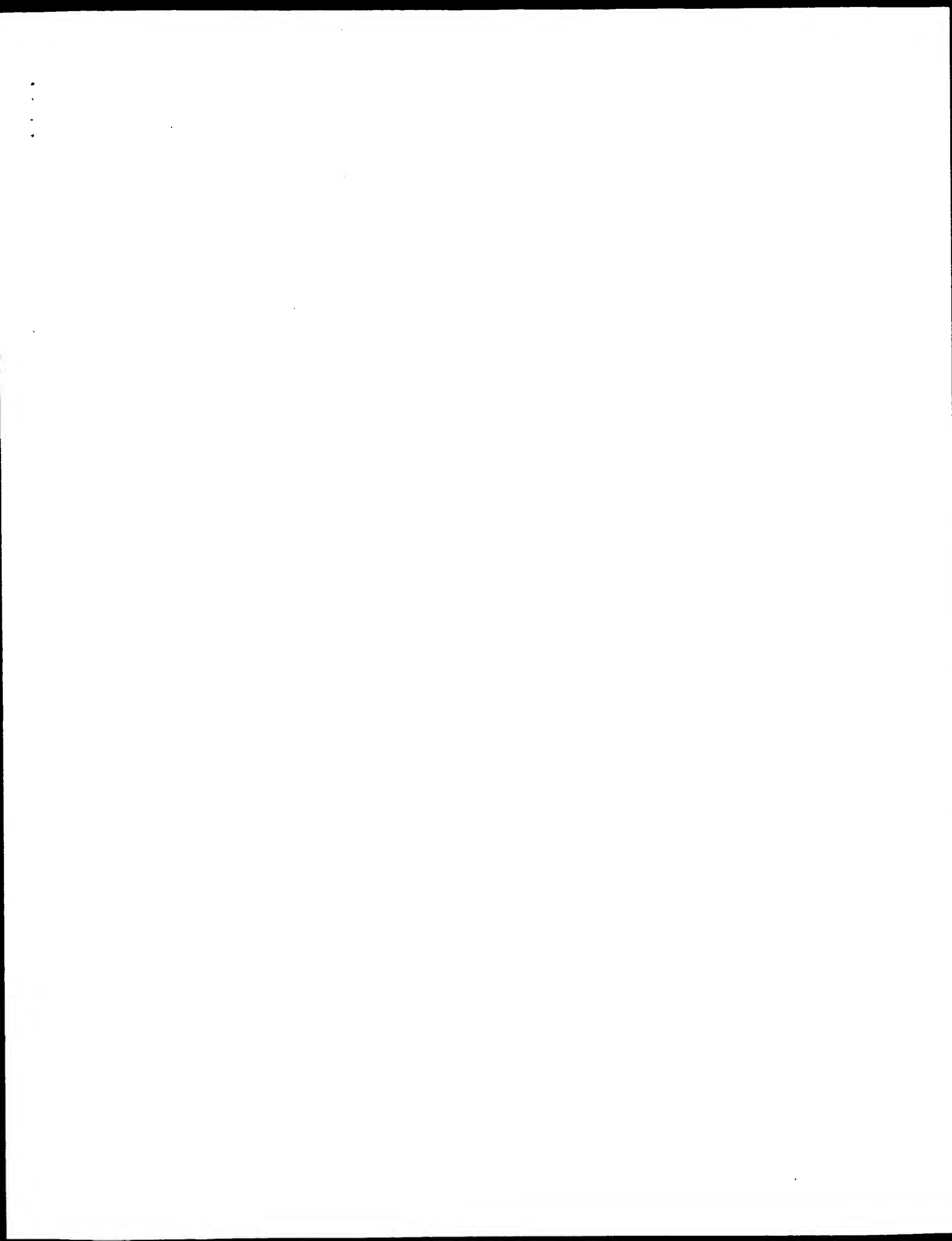
OY 1 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTTGADVREHIEPVLNRYGLPINORFIIV 60  
DB 4 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTTGADVREHIEPVLNRYGLPINORFIIV 63  
OY 61 ELSNHAELSVTLADVTNAVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAFG 120  
DB 64 ELSNHAELSVTLADVTNAVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAFG 123  
OY 121 GNYDRLEQLAGNLENIELGNPLEEASISALYYSTGTQDPLTARSFIICMISEAR 180  
DB 124 GNYDRLEQLAGNLENIELGNPLEEASISALYYSTGTQDPLTARSFIICMISEAR 183  
OY 181 FOYTEGEMRTIRIRNRRS 198  
DB 184 FOYTEGEMRTIRIRNRRS 201

Search completed: September 16, 2003, 11:51:46

Tue Sep 16 12:27:48 2003

us-10-083-336a-3.ra1

Job time : 11.2796 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:49:23 ; Search time 20.4768 Seconds  
(without alignments)  
1439.723 Million cell updates/sec

Title: US-10-083-336A-3

Perfect score: 1020  
Sequence: 1 IFPKOYPIINFTTAGATVQS.....ARFOYIEGEMRTIRYNRRS 198

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 556269 segs, 148893369 residues

Total number of hits satisfying chosen parameters: 556269

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubppa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubppa/PCF\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubppa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubppa/PCFUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubppa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubppa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubppa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubppa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubppa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1020	100.0	267	12 US-10-127-890-1	Sequence 1, Appli
2	1010	99.0	267	12 US-10-282-935-1	Sequence 1, Appli
3	342	33.5	247	10 US-09-792-793A-39	Sequence 39, Appli
4	342	33.5	247	10 US-10-127-890-6	Sequence 6, Appli
5	342	33.5	289	12 US-10-280-679B-4	Sequence 4, Appli
6	323.5	31.7	251	12 US-10-282-935-3	Sequence 3, Appli
7	312	30.6	247	10 US-09-792-793A-34	Sequence 34, Appli
8	279	27.4	263	12 US-10-127-890-4	Sequence 4, Appli
9	277	27.2	263	12 US-10-127-890-7	Sequence 7, Appli
10	276	27.1	252	9 US-09-347-064-2	Sequence 2, Appli
11	276	27.1	248	12 US-10-127-890-8	Sequence 8, Appli
12	267.5	26.2	248	12 US-10-127-890-5	Sequence 5, Appli
13	246.5	24.2	251	12 US-10-127-890-110	Sequence 110, App
14	246.5	24.2	251	12 US-10-127-890-111	Sequence 111, App
15	245.5	24.1	251	12 US-10-127-890-101	Sequence 101, App

16	244.5	24.0	251	12 US-10-127-890-99	Sequence 99, Appli
17	244.5	24.0	251	12 US-10-127-890-107	Sequence 107, App
18	243.5	23.9	251	12 US-10-127-890-100	Sequence 100, App
19	243.5	23.9	251	12 US-10-127-890-106	Sequence 106, App
20	242.5	23.8	251	9 US-09-765-527-247	Sequence 247, App
21	242.5	23.8	251	12 US-10-127-890-2	Sequence 2, Appli
22	242.5	23.8	251	12 US-10-127-890-102	Sequence 102, App
23	242.5	23.8	251	12 US-10-127-890-104	Sequence 104, App
24	242.5	23.8	251	12 US-10-127-890-105	Sequence 105, App
25	241.5	23.7	293	9 US-09-765-527-259	Sequence 259, App
26	241.5	23.7	293	9 US-09-765-527-253	Sequence 253, App
27	241.5	23.7	309	9 US-09-765-527-253	Sequence 253, App
28	241.5	23.7	309	9 US-09-765-527-253	Sequence 253, App
29	241.5	23.7	309	9 US-09-765-527-253	Sequence 253, App
30	240.5	23.6	251	12 US-10-127-890-108	Sequence 108, App
31	182	17.8	263	10 US-09-978-274A-4	Sequence 4, Appli
32	182	17.8	314	10 US-09-978-274A-2	Sequence 2, Appli
33	180.5	17.7	261	12 US-10-127-890-9	Sequence 9, Appli
34	176	17.3	250	10 US-09-792-793A-36	Sequence 36, Appli
35	176	17.3	250	12 US-10-127-890-8	Sequence 8, Appli
36	135	13.2	280	12 US-10-127-890-3	Sequence 3, Appli
37	135	13.2	280	12 US-10-127-890-127	Sequence 127, App
38	135	13.2	280	12 US-10-127-890-128	Sequence 128, App
39	135	13.2	280	12 US-10-127-890-129	Sequence 129, App
40	126	12.4	330	10 US-09-792-793A-82	Sequence 82, Appli
41	125	12.3	332	10 US-09-792-793A-73	Sequence 73, Appli
42	124	12.2	254	10 US-09-792-793A-85	Sequence 85, Appli
43	124	12.2	259	12 US-10-127-890-10	Sequence 35, Appli
44	124	12.2	275	10 US-09-792-793A-35	Sequence 79, Appli
45	124	12.2	337	10 US-09-792-793A-79	

## ALIGNMENTS

RESULT 1  
US-10-127-890-1  
Sequence 1, Application US/10127890  
Publication No. US2003016196A1

GENERAL INFORMATION:  
APPLICANT: Bectel, Marc D.  
Carroll, Stephen F.  
Studulka, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/03348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-127-890-1

Query Match 100.0%; Score 1020; DB 12; Length 267;  
Best Local Similarity 100.0%; Pred. No. 7e-109;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFPKOYPLINFTAGATVOSTYNTFIRAVGRRLTGGADVREIPLVLRVGLPINORFIIV 60  
DB 1 IFPKOYPLINFTAGATVOSTYNTFIRAVGRRLTGGADVREIPLVLRVGLPINORFIIV 60  
OY 61 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 120  
DB 61 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 120  
OY 121 GNYRLQAGNLRNENIELGNPLEAISALYYSTGGTOLPTLARSFTICIMISEAR 180  
DB 121 GNYRLQAGNLRNENIELGNPLEAISALYYSTGGTOLPTLARSFTICIMISEAR 180  
OY 181 FOYIEGEMRTIRIRYRRS 198  
DB 181 FOYIEGEMRTIRIRYRRS 198

RESULT 2  
US-10-282-935-1  
Sequence 1, Application US/10282935  
Publication No. US20030143193A1  
GENERAL INFORMATION:  
APPLICANT: VITERA, ELLEN S.  
APPLICANT: GHETTE, VICTOR F.  
APPLICANT: SMALLSHAW, JOAN  
APPLICANT: BALUNA, ROXANA G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF  
FILE REFERENCE: US20030143193A1  
CURRENT APPLICATION NUMBER: US/10/282,935  
CURRENT FILING DATE: 2002-10-29  
PRIOR APPLICATION NUMBER: 09/538,873  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: 60/126,826  
PRIOR FILING DATE: 1999-03-30  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 1  
LENGTH: 267  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-282-935-1

Query Match 99.0%; Score 1010; DB 12; Length 267;  
Best Local Similarity 100.0%; Pred. No. 9.9e-108;  
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PKOYPLINFTAGATVOSTYNTFIRAVGRRLTGGADVREIPLVLRVGLPINORFIIV 62  
DB 3 PKOYPLINFTAGATVOSTYNTFIRAVGRRLTGGADVREIPLVLRVGLPINORFIIV 62  
OY 63 SNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 122  
DB 63 SNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 122  
OY 123 YDRLEQAGNLRNENIELGNPLEAISALYYSTGGTOLPTLARSFTICIMISEAR 182  
DB 123 YDRLEQAGNLRNENIELGNPLEAISALYYSTGGTOLPTLARSFTICIMISEAR 182  
OY 183 YIEGEMRTIRIRYRRS 198  
DB 183 YIEGEMRTIRIRYRRS 198

RESULT 3  
US-09-792-793A-39  
Sequence 39, Application US/09792793A  
Patent No. US20020168370A1  
GENERAL INFORMATION:  
APPLICANT: McDonald, John R.  
APPLICANT: Coggin, Philip  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
FILE REFERENCE: 25020-601D  
CURRENT APPLICATION NUMBER: US/09/792,793A  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 39  
LENGTH: 247  
TYPE: PRT  
ORGANISM: Trichosanthens kirilowii  
US-09-792-793A-39

Query Match 33.5%; Score 342; DB 10; Length 247;  
Best Local Similarity 39.5%; Pred. No. 5.6e-31;  
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;

OY 9 INFTAGATVOSTYNTFIRAVGRRLTGGADVREIPLVLRVGLPINORFIIVLSNHAEL 68  
DB 2 VSFRLSGATVSSGVTISNLRKALPNERKL-IDIPPL-RSSLPSQRYALHLTNVADE 58  
OY 69 SVTLADVTNAVYVGRAGNSAYFFHPDNOEDA-ETHLFTDVONRYTFAGNYDRLE 127  
DB 59 TISVALDVNTVIMGRAGDTSYFF--NEASATEAKYVFKDAMRKVTLPSGNYERIQ 115  
OY 128 OLGNLRNENIELGNPLEAISALYYSTGGTOLPTLARSFTICIMISEARFOYIGE 187  
DB 116 TAAGKIRENIPDLIPALDSAITLFFYNN--SAASALWLIQSTSEARVFIQO 170  
OY 188 MRTIR 192  
DB 171 IGKRV 175

RESULT 4  
US-10-127-890-6  
Sequence 6, Application US/10127890  
Publication No. US20030166196A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studinka, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago

STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-May-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-May-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-May-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-Dec-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-Jun-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-Nov-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-9155  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 247 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-127-890-6  
Query Match 33.5%; Score 342; DB 12; Length 247;  
Best Local Similarity 39.5%; Pred. No. 5.6e-31;  
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
QY 9 INFTTACATVOSTNFRRAVGRLLTGADVRHEIPVLPNRYGLPINORFLVELSNHAEL 68  
DB 2 VSFRLSGATSSSYGVFISNLKALPNERKL-YDIPPL-RSSLPGSGORALHILHTNYADE 58  
QY 69 SVTLALDVTNAYVYGRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRRTFAFGNDRL 127  
DB 59 TISVAIDVTNAYVYGRAGDTSYFF--NEASATEAKYFKDMRKVTLPYSGNRYRLQ 115  
QY 128 QLAGNLRNIEIENGPLEAISALYYSTGTOLPTLARSFICICMISEARFYEGE 167  
DB 116 TAAGKIRENIPGLPALDSATITLTFYNNAN-----SAASALMWLIQSTSEARKYFTEQ 170  
QY 188 MTRRI 192  
DB 171 IGRKV 175

RESULT 5  
US-10-280-679B-4  
Sequence 4, Application US/10280679B  
Publication No. US20030150019A1  
GENERAL INFORMATION:  
APPLICANT: Large Scale Biology Corporation  
TITLE OF INVENTION: Monopartite RNA Virus Transformation Vectors  
FILE REFERENCE: LSBC-0109-US03  
CURRENT APPLICATION NUMBER: US/10/280,679B  
CURRENT FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: 09/557,941  
PRIOR FILING DATE: 2000-04-24  
PRIOR APPLICATION NUMBER: 08/484,341  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: 07/923,692  
PRIOR FILING DATE: 1992-07-31  
PRIOR APPLICATION NUMBER: 07/600,244  
PRIOR FILING DATE: 1990-10-22  
PRIOR APPLICATION NUMBER: 07/641,617  
PRIOR FILING DATE: 1991-01-16  
PRIOR APPLICATION NUMBER: 07/737,899  
PRIOR FILING DATE: 1991-07-26  
PRIOR APPLICATION NUMBER: 07/739,143  
PRIOR FILING DATE: 1991-08-01  
PRIOR APPLICATION NUMBER: 07/310,881  
PRIOR FILING DATE: 1989-02-17  
PRIOR APPLICATION NUMBER: 07/160,766  
PRIOR FILING DATE: 1988-02-26  
PRIOR APPLICATION NUMBER: 07/160,771  
PRIOR FILING DATE: 1988-02-26  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 289  
TYPE: PRT  
ORGANISM: Chinese cucumber protein alpha-trichosanthin  
US-10-280-679B-4

Query Match 33.5%; Score 342; DB 12; Length 289;  
Best Local Similarity 39.5%; Pred. No. 7e-31;  
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;

QY 9 INFTTACATVOSTNFRRAVGRLLTGADVRHEIPVLPNRYGLPINORFLVELSNHAEL 68  
DB 25 VSFRLSGATSSSYGVFISNLKALPNERKL-YDIPPL-RSSLPGSGORALHILHTNYADE 81  
QY 69 SVTLALDVTNAYVYGRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRRTFAFGNDRL 127  
DB 82 TISVAIDVTNAYVYGRAGDTSYFF--NEASATEAKYFKDMRKVTLPYSGNRYRLQ 138  
QY 128 QLAGNLRNIEIENGPLEAISALYYSTGTOLPTLARSFICICMISEARFYEGE 187  
DB 139 TAAGKIRENIPGLPALDSATITLTFYNNAN-----SAASALMWLIQSTSEARKYFTEQ 193  
QY 188 MTRRI 192  
DB 194 IGRKV 198

RESULT 6  
US-10-282-935-3  
Sequence 3, Application US/10282935  
Publication No. US20030143193A1  
GENERAL INFORMATION:  
APPLICANT: VITETTA, ELLEN S.  
APPLICANT: GHETTE, VICTOR F.  
APPLICANT: SWALSHAW, JOAN  
APPLICANT: BALUNA, ROXANA G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF  
FILE REFERENCE: QTSD:88405  
CURRENT APPLICATION NUMBER: US/10/282,935  
CURRENT FILING DATE: 2002-10-29  
PRIOR APPLICATION NUMBER: 09/538,873  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: 60/126,826  
PRIOR FILING DATE: 1999-03-30  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 251  
TYPE: PRT

ORGANISM: Abrus precatorius  
US-10-282-935-3

Query Match 31.7%; Score 323.5; DB 12; Length 251;  
Best Local Similarity 42.2%; Pred. No. 7.7e-29;  
Matches 79; Conservative 28; Mismatches 67; Indels 13; Gaps 5;

QY 9 INFTAGATVOSTYNTFIRAVRGLTTGADVRHEIPVLPNRVGLPINORFIVELSNHAEI 68  
DB 5 IKSTEGATOSYKOFIEALRELRG--LIHDIPVLPDPTLOERNRYIVELSNSDTE 62  
QY 69 SVTLADVTNAYVYVAGNSAYEFH--PDNDAEATLFTDVOYNTYFAFGNVDRL 126  
DB 63 SIEGIDVTNAYVYVAGNSAYEFH--PDNDAEATLFTDVOYNTYFAFGNVDRL 116  
QY 127 EOLAGNRENIENGPLEEASALYYSTGTQPLTARSFICIQMISEARFOYLEG 186  
DB 117 ERMAHQROQIPGLALTLTGIS--FERSGNDNEKARTLIVLIQVMAARFYISN 173  
QY 187 EMTRIR 193  
DB 174 RVRSIQ 180

RESULT 7  
US-09-792-793A-34  
Sequence 34, Application US/09792793A  
Patent No. US20020168370A1

GENERAL INFORMATION:  
APPLICANT: McDonald, John R.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND  
FILE REFERENCE: 25020-601D  
CURRENT APPLICATION NUMBER: US/09/792.793A  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 34  
LENGTH: 247  
TYPE: PRT  
ORGANISM: Bryonia dioica  
US-09-792-793A-34

Query Match 30.6%; Score 312; DB 10; Length 247;  
Best Local Similarity 37.3%; Pred. No. 1.6e-27;  
Matches 69; Conservative 46; Mismatches 58; Indels 12; Gaps 5;

QY 9 INFTAGATVOSTYNTFIRAVRGLTTGADVRHEIPVLPNRVGLPINORFIVELSNHAEI 68  
DB 2 VSFRLSGATTSYGVFIKNLREALPYERKV-YNIPLT--RSSISGSGRYTLHLTNVADE 58  
QY 69 SVTLADVTNAYVYVAGNSAYEFH--PDNDAEATLFTDVOYNTYFAFGNVDRL 127  
DB 59 TISVAVDTNAYVYVAGNSAYEFH--PDNDAEATLFTDVOYNTYFAFGNVDRL 115  
QY 128 OLAGNRENIENGPLEEASALYYSTGTQPLTARSFICIQMISEARFOYLEG 187  
DB 116 TAAKIRENIPGLALTLTGIS--FERSGNDNEKARTLIVLIQVMAARFYISN 173  
QY 188 MTRIR 192  
DB 171 IGRV 175

RESULT 8  
US-10-127-890-4  
Sequence 4, Application US/10127890  
Publication No. US20030166196A1

GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnicka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/054,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70. P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-127-890-4  
Query Match 27.4%; Score 279; DB 12; Length 263;  
Best Local Similarity 35.3%; Pred. No. 1.1e-23;  
Matches 66; Conservative 42; Mismatches 67; Indels 12; Gaps 5;

QY 9 INFTAGATVOSTYNTFIRAVRGLTTGADVRHEIPVLPNRVGLPINORFIVELSNHAEI 68  
DB 2 VNFDSLATPKTKTLEDEFRATLPFSHKV-YDIPLYSTIS--DSRRFILLDTLSYAVE 58  
QY 69 SVTLADVTNAYVYVAGNSAYEFH--PDNDAEATLFTDVOYNTYFAFGNVDRL 128  
DB 59 TISVAVDTNAYVYVAGNSAYEFH--PDNDAEATLFTDVOYNTYFAFGNVDRL 114  
QY 129 LAGNRENIENGPLEEASALYYSTGTQPLTARSFICIQMISEARFOYLEG 188  
DB 115 AAHKIRENIDGLPALSSATITLFTYNA-----QSAFSLVLIQVTAARFYISN 169  
QY 189 RTIRYN 195  
DB 170 AKYVATN 176

RESULT 9  
US-10-127-890-7  
Sequence 7, Application US/10127890







? APPLICATION NUMBER: US/08/5446.360  
 ? FILING DATE: 13-MAY-1998  
 ? APPLICATION NUMBER: PCT/US94/05344  
 ? FILING DATE: 12-MAY-1994  
 ? APPLICATION NUMBER: US 08/064,691  
 ? FILING DATE: 12-MAY-1993  
 ? APPLICATION NUMBER: US 07/988,430  
 ? FILING DATE: 09-DEC-1992  
 ? APPLICATION NUMBER: US 07/901,707  
 ? FILING DATE: 19-JUN-1992  
 ? APPLICATION NUMBER: US 07/787,567  
 ? FILING DATE: 04-NOV-1991

1 APPLICATION NUMBER: US/08/6446,366  
2 FILING DATE: 13-MAY-1996  
3 APPLICATION NUMBER: PCT/US94/05340  
4 FILING DATE: 12-MAY-1994  
5 APPLICATION NUMBER: US 08/064,691  
6 FILING DATE: 12-MAY-1993  
7 APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 101:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 101:  
US-10-127-890-101

Query Match 24.1%; Score 245.5; DB 12; Length 251;  
Best Local Similarity 37.1%; Pred. No. 7.2e-20;  
Matches 72; Conservative 31; Mismatches 70; Indels 21; Gaps 9;

QY 9 INFTAGATVOSTYTPIRAVRGRLTGGADVREIPLVLPINQRFILVETSNHAEI 68  
DB 5 VSFSTKGATYITVYNFLNELRYKLKPEGN-SHGIPLRKKADDP-GKAFVLVALSNDNGQ 62  
69 SVTLADVTNAYVGYRAGNSAVFPH--PDNQDEAETHTLFTD-VQNRVTFAGGNYDR 125  
DB 63 LAEIAIDVTSVYVGQVNRBSYFPKADP----AAVEGLFKNTIKTR--LHFGGSYPS 115  
QY 126 LEQLAGN--LRNIEIENGPLEEALSAIYISTGQPLPTARSFTICIQMISEARQY 183  
DB 116 LE---GEKAYRETTDGLIEPLRIGIKIDENAIIDNYKPLEIASLLVVIQMVSEARPTF 172  
QY 184 IEGEMRT---RIR 193  
DB 173 IENQIRNNFQQRIR 186

Search completed: September 16, 2003, 12:09:56  
Job time : 21.4768 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:40:32; Search time 12.8414 Seconds

(without alignments)  
1482.817 Million cell updates/sec

Title: US-10-083-336A-3

Perfect score: 1020

Sequence: 1 IFPKOYPIINFETAGATVQS.....ARFOYIEGEMRTIRNRRS 198

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	100.0	576	1	RLCSD
2	930.5	91.2	564	1	ricin D precursor
3	353.5	34.7	528	2	agglutinin precursor
4	353.5	34.7	562	2	ablin-d precursor
5	345	33.8	527	2	ablin-c precursor
6	342	33.5	289	1	ablin-D precursor
7	338	33.1	247	2	RNA N-glycosidase
8	338	33.1	247	2	karasurin - Mongol
9	338	33.1	289	2	karasurin-B - Tric
10	336.5	33.0	251	2	karasurin C - Tric
11	329.5	32.3	528	1	abrin (clone 7.2)
12	307.5	30.1	278	2	ablin-a precursor
13	300.5	28.5	250	2	beta-luifin - smoo
14	291	28.0	570	2	luifin-b - smooch
15	286	27.9	254	2	agglutinin I precu
16	284.5	27.9	277	2	mistletoe lectin I
17	279	27.4	286	2	RNA N-glycosidase
18	278	27.3	245	2	RNA N-glycosidase
19	277	27.2	286	1	RNA N-glycosidase
20	272	26.7	286	1	RNA N-glycosidase
21	242.5	23.8	316	2	RNA N-glycosidase
22	197.5	19.4	294	2	RNA N-glycosidase
23	182	17.8	313	2	RNA N-glycosidase
24	180.5	17.7	261	2	RNA N-glycosidase
25	178	17.5	272	2	antiviral protein
26	156	15.3	278	2	RNA N-glycosidase
27	149.5	14.7	289	2	betavulgin - beet
28	140	13.7	280	1	RNA N-glycosidase
29	137	13.4	275	2	RNA N-glycosidase

30	133	13.0	281	2	B38664	30K ribosome inact
31	132	12.9	280	2	JC5848	protein synthesis
32	131	12.8	253	2	S28542	RNA N-glycosidase
33	127	12.5	253	2	S28539	RNA N-glycosidase
34	127	12.5	253	2	S29931	RNA N-glycosidase
35	124	12.2	283	2	S05205	RNA N-glycosidase
36	123.5	12.1	293	2	S17519	RNA N-glycosidase
37	123	12.1	253	2	S28541	RNA N-glycosidase
38	122	12.0	310	2	S46232	ribosome-inactivat
39	121	11.9	292	1	RLQHS2	RNA N-glycosidase
40	115	11.3	253	2	AS8923	RNA N-glycosidase
41	111	10.9	319	2	I54695	shiga-like toxin I
42	111	10.9	319	2	E90779	shiga toxin 2 subu
43	111	10.9	319	2	C85640	Shiga toxin 2 subu
44	110.5	10.8	318	2	S01032	Shiga-like toxin I
45	110	10.8	236	2	S17932	RNA N-glycosidase

## ALIGNMENTS

## RESULT 1

RLCSD  
ricin D precursor - castor bean  
N:Contains: RNA N-glycosidase (EC 3.2.2.22)  
C:Species: Ricinus communis (castor bean)  
C:Date: 31-Dec-1993 #sequence-revision 31-Dec-1993 #text-change 16-Jul-1999  
C:Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903  
R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.F.  
Nucleic Acids Res. 13, 8019-8033, 1985  
A:Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.  
A:Reference number: A24041; MUID:86067214; PMID:2999712  
A:Accession: A24041  
A:Molecule type: DNA  
A:Residues: 1-576 <HAL>  
A:Cross-references: GB:X03179; NID:g21082; PIDN:CAA26939.1; PID:g21083  
R:Riegger, J.W.; Roberts, L.M.  
Plant Mol. Biol. 18, 515-525, 1992  
A:Title: The lectin gene family of Ricinus communis: cloning of a functional ricin ge  
A:Reference number: S20513; MUID:92163016; PMID:1371405  
A:Accession: S20513  
A:Molecule type: DNA  
A:Residues: 1-576 <TRE>  
A:Cross-references: EMBL:X52908; NID:g21084; PIDN:CAA37095.1; PID:g21085  
R:Ramb, F.I.; Roberts, L.M.; Lord, J.M.  
Eur. J. Biochem. 148, 265-270, 1985  
A:Title: Nucleotide sequence of cloned cDNA coding for preproricin.  
A:Reference number: A24614; MUID:85179479; PMID:3838723  
A:Accession: A24614  
A:Molecule type: mRNA  
A:Residues: 12-75, 'D', '77-550, 'R', '552-576 <LAM>  
A:Cross-references: GB:X02388; NID:g21077; PIDN:CAA26230.1; PID:g21078  
R:Yoshitake, S.; Funatsu, G.; Funatsu, M.  
Agric. Biol. Chem. 42, 1267-1274, 1978  
A:Title: Isolation and sequences of peptic peptides, and the complete sequence of ile  
A:Reference number: A03372  
A:Accession: A03372  
A:Molecule type: protein  
A:Residues: 36-97, 'Q', '99-109, 'S', '111-269, 'D', '272-283, 'L', '285-288, 290-302 <YOS>  
A:Note: This paper cites the others in the series providing experimental details for  
R:Ataki, T.; Funatsu, G.  
FEBS Lett. 191, 121-124, 1985  
A:Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptop  
A:Reference number: A24010  
A:Accession: A24010  
A:Molecule type: protein  
A:Residues: 315-383, 'PS', '386-576 <ARA>  
R:Funatsu, G.; Kimura, M.; Funatsu, M.  
Agric. Biol. Chem. 43, 2221-2224, 1979  
A:Title: Primary structure of Ala chain of ricin D.  
A:Reference number: A03374  
A:Accession: A03374  
A:Molecule type: protein

A:Residues: 315-335, 'N', 337-342, 'NH', 345-362, 364-383, 'PS', 386-399, 'T', 401, 'D', 403, 'E', 405, 527, 'E', 529-564, 'W', 566, 'H', 567-570, 'LI', 573-574, 'F', <FNU>  
 A:Note: This paper, one of a series, summarizes the experimental details for the determination of the complete amino acid sequence of the B-chain of the Ricinus communis agglutinin (RiAgg), M.P., Kim, Y., Robertus, J.D.  
 A:Title: Site-directed mutagenesis of ricin A-chain and implications for the mechanism of action of ricin A-chain  
 A:Reference number: A48237; PMID:91352006; PMID:1881883  
 A:Contents: annotation: active site  
 R:Robertus, J.D.  
 A:Title: Structure of ricin B-chain at 2.5 angstrom resolution.  
 A:Reference number: A48238; PMID:91352005; PMID:1881882  
 A:Contents: annotation: X-ray crystallography, 2.5 angstroms  
 R:Katzin, B.J.; Collins, E.J.; Robertus, J.D.  
 A:Title: Structure of ricin A-chain at 2.5 angstroms.  
 A:Reference number: A48239; PMID:91352004; PMID:1881881  
 A:Contents: annotation: X-ray crystallography, 2.5 angstroms  
 C:Comment: The functional molecule is a disulfide-linked dimer of A and B chains, which onto the cell of the A chain. B chains are also responsible for cell agglutination (lectin C:Comment: This protein is cytotoxic and very poisonous to animals.  
 C:Superfamily: ricin; RNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed  
 F:1-35/Domain: signal sequence #status predicted <SIG>  
 F:36-302/Product: ricin D chain A #status experimental <ACH>  
 F:46-293/Domain: RNA N-glycosidase homology <RNG>  
 F:315-373/Product: ricin D chain B #status experimental <ECH>  
 F:331-373-374-414-417-455-462-497-501-540-543-576/Region: 40-residue repeats  
 F:449-449/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:115-158-243-244/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:212/Active site: Arg #status predicted  
 F:215/Active site: Arg #status predicted  
 F:294-318-334-353-377-394-465-478-504-521/Disulfide bonds: #status experimental  
 F:336-349-360/Binding site: N-acetylgalactosamine (Asp, Glu, Asn) #status experimental  
 F:548-569/Binding site: N-acetylgalactosamine (Asp, Asn) #status experimental

Query Match 100.0%; Score 1020; DB 1; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 4, 4e-84;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFPKOPIINFTAGATVOSTNFRIRAVGRRLTGADVHREIPVLRVGLPINORITLV 60  
 DB 36 IFPKOPIINFTAGATVOSTNFRIRAVGRRLTGADVHREIPVLRVGLPINORITLV 95  
 QY 61 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPNOEDAEATLHFTDQNRYPFAG 120  
 DB 96 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPNOEDAEATLHFTDQNRYPFAG 155  
 QY 121 GNTDRLEOLAGNIRENIEGNGPLEAISALYYSTGQTOLPTLARSFTICMISEAR 180  
 DB 156 GNTDRLEOLAGNIRENIEGNGPLEAISALYYSTGQTOLPTLARSFTICMISEAR 215  
 QY 181 FOYIEGEMRTIRYNNRS 198  
 DB 216 FOYIEGEMRTIRYNNRS 233

RESULT 2  
 RLCASG  
 agglutinin precursor - castor bean  
 N:Contains: Ricinus communis (castor bean)  
 C:Species: Ricinus communis (castor bean)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
 C:Accession: A24261; A24210  
 R:Roberts, L.M.; Lamb, I.F.; Paparin, D.J.C.; Lord, J.M.  
 J. Biol. Chem. 260, 15682-15686, 1985  
 A:Title: The primary sequence of Ricinus communis agglutinin. Comparison with ricin.  
 A:Reference number: A24261; PMID:86059449; PMID:2991310  
 A:Molecule type: mRNA  
 A:Residues: 1-564 <ROB>  
 A:Cross-references: GB:M12089; NID:q169700; PIDN:AAA33868.1; PID:q169701

R:Araki, T.; Yoshioke, Y.; Funatsu, G.  
 Biochim. Biophys. Acta 872, 277-285, 1986  
 A:Title: The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin (RiAgg), M.P., Kim, Y., Robertus, J.D.  
 A:Reference number: A24210  
 A:Accession: A24210  
 A:Molecule type: protein  
 A:Residues: 303-325, 'F', 327-330, 'T', 332-361, 'D', 363-373, 'G', 375-403, 'T', 405-551, 'V', 5  
 A:Comment: This protein has strong agglutinating activity and weak cytotoxicity compa  
 C:Superfamily: ricin; RNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; s  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-290/Product: agglutinin chain A #status predicted <ACH>  
 F:35-281/Domain: RNA N-glycosidase homology <RNG>  
 F:303-564/Product: agglutinin chain B #status experimental <ECH>  
 F:319-361-362-402-405-443-450-485-489-528-531-564/Region: 40-residue repeats  
 F:34-259/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:104-147-231-232/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:200-203/Active site: Glu, Arg #status predicted  
 F:282-306-322-341-365-382-453-466-492-509/Disulfide bonds: #status predicted  
 F:324-337-348/Binding site: N-acetylgalactosamine (Asp, Glu, Asn) #status predicted  
 F:397-437/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:536-557/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 91.2%; Score 930.5; DB 1; Length 564;  
 Best Local Similarity 91.9%; Pred. No. 5e-76;  
 Matches 182; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 1 IFPKOPIINFTAGATVOSTNFRIRAVGRRLTGADVHREIPVLRVGLPINORITLV 60  
 DB 25 IFPKOPIINFTAGATVOSTNFRIRAVGRRLTGADVHREIPVLRVGLPINORITLV 84  
 QY 61 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPNOEDAEATLHFTDQNRYPFAG 120  
 DB 85 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPNOEDAEATLHFTDQNRYPFAG 144  
 QY 121 GNTDRLEOLAGNIRENIEGNGPLEAISALYYSTGQTOLPTLARSFTICMISEAR 180  
 DB 145 GNTDRLEOLAGNIRENIEGNGPLEAISALYYSTGQTOLPTLARSFTICMISEAR 203  
 QY 181 FOYIEGEMRTIRYNNRS 198  
 DB 204 FOYIEGEMRTIRYNNRS 221

RESULT 3  
 S32431  
 abrin-d precursor - Indian licorice (fragment)  
 N:Contains: Ricinus communis (Indian licorice)  
 C:Species: Ricinus communis (Indian licorice)  
 C:Date: 30-Sep-1993 #sequence\_revision 01-Aug-1997 #text\_change 01-Aug-1997  
 C:Accession: S32431; S34408  
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
 J. Mol. Biol. 229, 263-267, 1993  
 A:Title: Primary structure of three distinct isoabrin determined by cDNA sequencing.  
 A:Reference number: S32429; PMID:93132798; PMID:8421313  
 A:Accession: S32431  
 A:Molecule type: mRNA  
 A:Residues: 1-528 <HUN>  
 A:Cross-references: GB:M98346  
 R:Hung, C.; Lee, M.; Lee, T.; Lin, J.  
 submitted to the EMBL Data Library, March 1993  
 A:Reference number: S34408  
 A:Accession: S34408  
 A:Molecule type: mRNA  
 A:Residues: 1-169, 'C', 171-320, 'V', 322-528 <HU2>  
 A:Cross-references: GB:M98346  
 C:Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactiva  
 The A and B chains are linked by a single disulfide bond, which is essential for tox  
 C:Superfamily: ricin; RNA N-glycosidase homology  
 C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin  
 F:1-251/Product: abrin-d chain A #status predicted <ACH>  
 F:7-246/Domain: RNA N-glycosidase homology <RNG>  
 F:261-528/Product: abrin-d chain B #status predicted <ECH>



C:Date: 30-Sep-1998 #sequence revision 26-Jan-1996 #text\_change 23-Mar-2001  
A:Accession: J70566; A36274; JCI1093; A36273; J70003  
R:Shaw, P.C.; Yung, M.H.; Zhu, R.H.; Ho, W.K.K.; Ng, T.B.; Yeung, H.W.  
Gene 97, 267-272, 1991  
A:Title: Cloning of trichosanthin cDNA and its expression in *Escherichia coli*.  
A:Reference number: J70566; MUID:9153657; PMID:1999291  
A:Accession: J70566  
A:Molecule type: mRNA  
A:Residues: 1-289 <SHA>  
A:Cross-references: GB:M4858; NID:g170536; PIDN:AAA34207.1; PID:g170537  
R:Chow, T.P.; Feldman, R.A.; Lovett, M.; Piatlak, M.  
J. Biol. Chem. 265, 8670-8674, 1990  
A:Title: Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type I rib  
A:Reference number: A36274; MUID:90256790; PMID:2341480  
A:Accession: A36274  
A:Molecule type: DNA  
A:Residues: 1-233,'T', 235-246,'M', 248-289 <CHO>  
A:Cross-references: GB:J05434; NID:g170534; PIDN:AAA34206.1; PID:g170535  
R:Zhang, H.G.; Wang, B.; Shao, P.Z.; Yang, X.R.  
Acta Genet. Sin. 21, 42-51, 1994  
A:Title: Cloning and DNA sequencing of the gene encoding Trichosanthin.  
A:Reference number: JCI1093; MUID:94271613; PMID:8003348  
A:Accession: JCI1093  
A:Molecule type: DNA  
A:Residues: 1-72,'V', 74-90,'S', 92-233,'T', 235-267,'D', 269-289 <ZHE>  
A:Cross-references: GB:S70176; NID:g9547148; PIDN:AA831048.1; PID:g9547149  
R:Collins, E.J.; Robertus, J.D.; Lopresti, M.; Stone, K.L.; Williams, K.R.; Wu, P.; Hwan  
J. Biol. Chem. 265, 8665-8669, 1990  
A:Title: Primary amino acid sequence of alpha-trichosanthin and molecular models for ab  
A:Reference number: A36273; MUID:90256789; PMID:2341399  
A:Accession: A36273  
A:Molecule type: Protein  
A:Residues: 24-270 <COL>  
R:Wang, Y.; Qian, R.Q.; Gu, Z.W.; Jin, S.W.; Zhang, L.Q.; Xia, Z.X.; Tian, G.Y.; Ni, C.Z.  
Pure Appl. Chem. 58, 789-798, 1986  
A:Title: Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application  
A:Reference number: J70003  
A:Accession: J70003  
A:Molecule type: Protein  
A:Residues: 24-56,'L', 58-59,'T', 61-71,'T', 73-81,85-86,'L', 88-92,'DAGLPNNAV', 93-142,'GL'  
A:Experimental source: tuber  
R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.  
submitted to the Brookhaven Protein Data Bank, July 1994  
A:Reference number: A67091; PDB:1MRJ  
A:Contents: annotation: X-ray crystallography, 1.6 angstroms, with adenine, residues 24-  
R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.  
submitted to the Brookhaven Protein Data Bank, July 1994  
A:Reference number: A67092; PDB:1MRK  
A:Contents: annotation: X-ray crystallography, 1.6 angstroms, with formycin, residues 24-  
R:Xiong, J.P.; Xia, Z.X.; Wang, Y.  
submitted to the Brookhaven Protein Data Bank, December 1994  
A:Reference number: A66711; PDB:1TCS  
A:Contents: annotation: X-ray crystallography, 1.7 angstroms, with NADPH, residues 24-27  
R:Xiong, J.P.; Xia, Z.X.; Wang, Y.  
Nat. Struct. Biol. 1, 695-700, 1994  
A:Title: Crystal structure of trichosanthin-NADPH complex at 1.7 Angstroms resolution re  
A:Reference number: A58622; MUID:95360714; PMID:7634073  
A:Contents: annotation: X-ray crystallography, 1.7 angstroms  
C:Comment: Alpha-trichosanthin has been used to induce abortions.  
C:Genetics:  
A:Gene: tcs  
C:Function:  
A:Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA there  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: abortifacient; glycosidase; hydrolase; root; toxin  
E:1-23/Domains: signal sequence #status predicted <SIG>  
F:24-270/Product: trichosanthin alpha #status experimental <MAT>  
F:271-266/Domains: rRNA N-glycosidase homology <RNO>  
F:271-289/Domains: carboxyl-terminal propeptide #status predicted <CTP>  
F:93,183,186/Active site: Tyr, Glu, Arg #status predicted

RESULT 7  
JU0393  
karasurin - Mongolian snake-gourd  
C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 12-Apr-1995  
C:Accession: JU0393; PS0163  
R:Toyokawa, S.; Takeda, T.; Kato, Y.; Wakabayashi, K.; Ogihara, Y.  
Chem. Pharm. Bull. 39, 1244-1249, 1991  
A:Title: The complete amino acid sequence of an abortifacient protein, karasurin  
A:Reference number: JU0393; MUID:92005921; PMID:1914000  
A:Accession: JU0393  
A:Molecule type: protein  
A:Residues: 1-247 <OY>  
A:Note: a sequence which lacks Ala-247 is also shown in this publication  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C:Keywords: abortifacient  
E:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match	33.1%	Score 338;	DB 2;	Length 247;
Best Local Similarity	40.0%;	Pred. No. 4.6e-23;		
Matches 74;	Conservative 46;	Mismatches 53;	Indels 12;	Gaps 5

```

QY      9 INFTAGATVQSYTNFTLRVRGRRLTGADVHEHPVLNRYGLP INQRFILVELSNAEI 68
      2 VSFRLSGATSSSYGVFISNKKRALPYEKRL-YDIPIL--RSTLPGSRVALIHILNTADE 58
Db
QY      69 SVTTLALDTNNVYVGYFRAGNSAYVEFHDPNDEDA-EATHTLEFTDVQNYRTEAFGNNYDRE 127
      59 TISVALDVTNNVYVGYRAGDTSYFF--NEASATEAKKYVEKDAKRKRYTLPGSGNYDRLQ 115
Db
QY      128 QLAGNLKRNIELGNQPLEEALISALYYSTGTQLPILARSLICIMISEARQYVEGE 187
      116 IAAQGRINRIPLGLPALDSATTTLEYYNAN----SAASALMVLIOSTSEAKRYKFLTEQQ 170
QY      188 MRRRI 192
      : 1:
Db      171 IGRV 175

```





A:Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein fr  
A:Reference number: J0202  
A:Accession: J0202  
A:Molecule type: protein  
A:Residues: 1-201,203-251 <RUN>  
A:Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have  
R:Evensen, G.; Mathiesen, A.; Sundan, A.  
J. Biol. Chem. 266, 6848-6852, 1991  
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.  
A:Reference number: A39761; MUID:91201329; PMID:2016300  
A:Accession: A39761  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 'E',2-251 <EVE>  
A:Cross-references: GB:X54872  
A:Note: residues 1-8 were derived from the synthesized primer  
R:Kimura, M.; Sumizawa, T.; Funatsu, G.  
Biosci. Biotechnol. Biochem. 57, 166-169, 1993  
A:Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic  
A:Reference number: J01398; MUID:93169023; PMID:7763422  
A:Contents: seeds  
A:Accession: J01398  
A:Molecule type: protein  
A:Residues: 261-347,'T',349-351,'A',353-357,'L',359-528 <RIM>  
A:Experimental source: seed  
R:Evensen, G.; Mathiesen, A.; Sundan, A.  
submitted to the EMBL data library, October 1990  
A:Description: Direct molecular cloning of two distinct abrin A-chains.  
A:Reference number: S14471  
A:Accession: S14472  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 'ME',2-2251 <EV2>  
A:Cross-references: EMBL:X54873; NID:q16090; PIDN:CA38655.1; PID:q16091  
R:Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.  
FEBS Lett. 309, 115-118, 1992  
A:Title: The complete primary structure of abrin-a B chain.  
A:Reference number: S24133; MUID:92371656; PMID:1505674  
A:Accession: S24133  
A:Molecule type: protein  
A:Residues: 262-297,'Y',299-426,'L',428-466,'P',468-482,'L',484-528 <CHE>  
R:Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.  
Eur. J. Biochem. 240, 564-569, 1996  
A:Title: Probing the domain structure of abrin-a by tryptic digestion.  
A:Reference number: S74110; MUID:97008945; PMID:8856055  
A:Accession: S74110  
A:Molecule type: protein  
A:Residues: 89-108;154-172 <LIN>  
A:Experimental source: seed  
A:Accession: S74111  
A:Molecule type: protein  
A:Residues: 262-276,'X',278-280;329-348;369-388;399-418 <LIM>  
A:Experimental source: seed  
A:Comment: Abrin-a is more toxic than ricin. The toxin consists of an A chain, which inh  
taining receptors on the cell surface. The A and B chains are linked by a single disulf  
C:Superfamily: ricin; rRNA N-glycosidase homology  
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglyutamic acid  
F:1-251/Product: abrin-a chain A #status experimental <ACH>  
F:7-246/Domain: rRNA N-glycosidase homology <RNG>  
F:261-528/Product: abrin-a chain B #status experimental <BCH>  
F:283-325;326-366;369-407;414-449;453-492;495-528/Region: 40-residue repeats  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:14,113,195,196/Binding site: substrate (Tyr, Glu, Asn) #status predicted  
F:164,167/Active site: Glu, Arg #status predicted  
F:247-269,286-305,329-346,417-430,456-473/Disulfide bonds: #status predicted  
F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
F:361,401/Binding site: carboxylate (Asn) (covalent) #status experimental  
F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

OY	9	INFTGGATVQSTVNTNFRFAVRGLTTGADVRRHEPVPINQRFLLVLSNHAEL	68
		1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	
Db	5	IKSTSEGATQSOSTKQFLEALRERLRCG--LIHIDPVLPDPTTIOQRRIKTYTVELSSDTE	62
OY	69	SVTLALDVTNAAVYGVYRAGNSAYFEH--PDNOEDAEATLHFTFDVONRYTFAFGNDRIL	126
		1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	
Db	63	SIEVGIDVTNAAVYVAYRAGTQSYFLRDPAPSSASD----YLFGTG--DOHSLPEYGTGYDL	116
OY	127	EQLAGNLRNEMIELNGNLEEAISALYYVSGNGQDLPFLARSFFICIDMISEARFOYIEG	186
		1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	
Db	117	ERMHQSRQOQPIGLQALTHGIS---FFRSGNDNEKARTLLVIITQVAEAAERFIISN	173
OY	187	EMFTIR 193	
		1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100	
Db	174	RVRYSIQ 180	

RESULT 12

S23519

beta-luffin - smooth loofah

C:Species: Luffa cylindrica (smooth loofah)

C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 20-Aug-1999

C:Accession: S23519; S23113

R:Kataoka, J.; Habuka, N.; Miyano, M.; Masuta, C.; Kolwal, A.

Plant Mol. Biol. 19, 887-893, 1992

A:Title: Nucleotide sequence of cDNA encoding beta-luffin, another ribosome-inactivating protein from Luffa cylindrica

A:Reference number: S23519; MUID:92353400; PMID:1643290

A:Accession: S23519

A:Molecule type: mRNA

A:Residues: 1-278 <KAT>

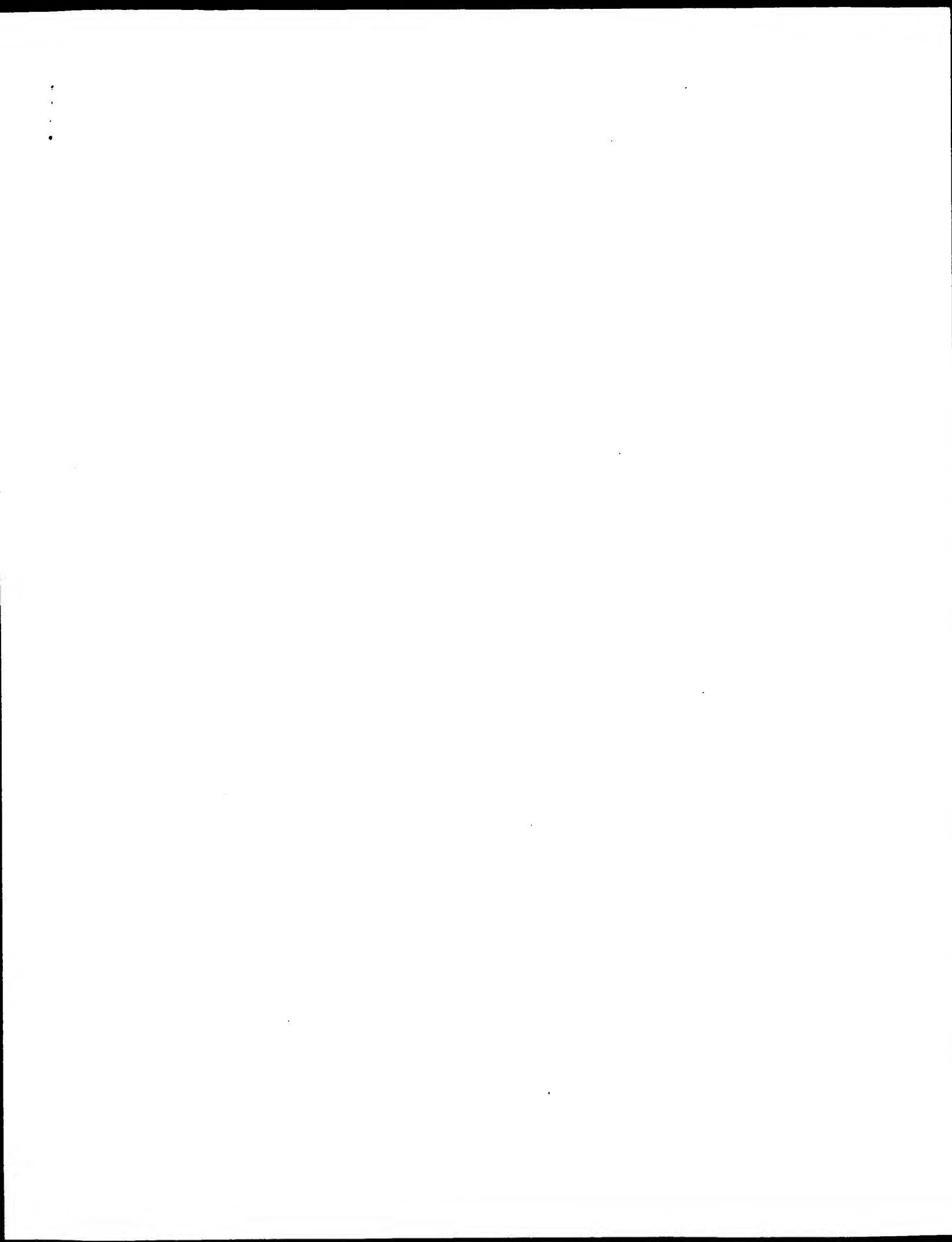
A:Cross-references: EMBL:X62372; NID:g19149; PIDN:CAA44230.1; PID:g19150

C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology

E:26-264/domain: rRNA N-glycosidase homology <RNG>

[illegible]





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:29:47 ; Search time 32.9711 seconds  
(Without alignments)  
953.195 Million cell updates/sec

Title: US-10-083-336a-3

Perfect score: 1020

Sequence: 1 IFPKOYPIINFTTAGATVQS.....AFQYIEGEMKRTIRYRRRS 198

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03:\*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	100.0	267	14 AAR37290	Ricin A chain. Un
2	1020	100.0	267	16 AAR63902	Ricin A-chain (RTA
3	1020	100.0	290	18 AAR25136	Ricin A-chain r1bo
4	1020	100.0	290	18 AAR21699	Ricin A-chain RIP.
5	1020	100.0	332	8 AAP70097	Ricin A. Escheric
6	1020	100.0	332	8 AAP70838	Sequence of Ricinu
7	1020	100.0	332	10 AAP95639	Ricin A encoded by
8	1020	100.0	554	16 AAR70827	Anti-catact immu
9	1020	100.0	562	10 AAP90079	Ricin D. Ricinus

10	1020	100.0	565	22 AAG78304	Modified castor be
11	1020	100.0	576	8 AAP70326	Sequence of Ricinu
12	1020	100.0	576	18 AAR25787	Castorbean ricin.
13	1020	100.0	576	20 AAY55892	Castor bean ricin
14	1020	100.0	576	21 AAV78592	Ricinus communis r
15	1020	100.0	576	22 AAG78301	Castor bean prepro
16	1020	100.0	576	22 AAG78302	Castor bean prepro
17	1017	99.7	565	6 AAP50166	Sequence of prepro
18	1017	99.7	565	22 AAG78300	Castor bean prepro
19	1016	99.6	200	9 AAR80164	Biosynthetic multi
20	1013	99.3	268	14 AAR39570	Sequence of ricin-
21	1013	99.3	574	8 AAP70325	Sequence of Ricinu
22	1012	99.2	565	7 AAP60240	Preproricin. Ricl
23	1011	99.1	574	10 AAP94793	DNA sequence of ri
24	1010	99.0	267	13 AAR30722	Ricin A from plicil
25	1010	99.0	267	21 AAB19265	Amino acid sequenc
26	1010	99.0	534	14 AAR39571	Sequence of G-FIT.
27	1009	98.9	332	11 AAR06554	Ricin A gene produ
28	993	97.4	267	16 AAR74176	Ricin A chain (RTA
29	987	96.8	267	14 AAR32430	Ricin A. Syntheti
30	930.5	91.2	540	18 AAR25143	Castor oil plant a
31	930.5	91.2	540	18 AAR21706	R. communis aglut
32	796.5	78.1	534	8 AAP70324	Sequence of Ricin
33	778	76.3	280	10 AAR95648	Ricin agglutinin A
34	342	33.5	247	16 AAR67359	Trichosanthin A
35	342	33.5	247	21 AAY69048	Amino acid sequenc
36	342	33.5	248	11 AAR07518	Synthetic alpha-tr
37	342	33.5	248	13 AAR25573	Mature alpha-trich
38	342	33.5	267	18 AAR25140	Trichosanthin anti
39	342	33.5	267	18 AAR21703	Trichosanthin (a r
40	342	33.5	289	11 AAR07514	Trichosanthin from
41	342	33.5	289	13 AAR25572	Trichosanthin prot
42	342	33.5	289	13 AAR29772	Trichosanthin prot
43	342	33.5	289	14 AAR32986	Encodes chinese cu
44	342	33.5	289	15 AAR51129	Alpha-trichosanthin
45	342	33.5	289	18 AAR10468	Chinese cucumber a

## ALIGNMENTS

RESULT 1	
AAAR37290	
ID	AAAR37290 standard; protein; 267 AA.
XX	
AC	AAAR37290:
XX	
DT	25-MAR-2003 (updated)
DT	09-JAN-2003 (updated)
DT	13-SEP-1993 (first entry)
XX	
DE	Ricin A chain.
XX	
KW	Type II ribosome-inactivating protein; type II RIP; gelonin;
KW	memoridin; immunocjugate; autoimmune disease; cell killing; toxin.
OS	Unidentified.
XX	
PN	WO9309130-A1.
XX	
PD	13-MAY-1993.
XX	
PF	04-NOV-1992: 92WC-US09487.
XX	
PR	04-NOV-1991: 91US-0787567.
PR	19-JUN-1992: 92US-0901707.
XX	
PA	(XOMA ) XOMA CORP.
XX	
PI	Berhard SL, Better MD, Carroll SF, Lane JA, Lei SP;
XX	WFI, 1993-167617/20.
XX	

PT Analogues of type I ribosome inactivating protein - useful as  
 PT cytotoxic agents, immuno toxins for treating auto immune diseases,  
 PT cancer, graft versus host disease and selective cell killing in-vivo  
 PS Claim 1; Page 92; 163pp; English.  
 XX  
 PS The invention covers analogues of Type I RIPS. Ricin is a Type II  
 CC RIP whose A chain is homologous to plant type I RIPS. The analogues  
 CC of the invention have a cysteine available for intermolecular  
 CC disulphide bonding at an amino acid position corresp. to a position  
 CC not naturally available for bonding; the cys residue is located in  
 CC the C-terminal region of the analogue between a position corresp. to  
 CC amino acid 251 and the C-terminus of ricin A chain. The analogues are  
 CC pref. joined via a disulphide linkage to a molecule which specifically  
 CC binds to a target cell, e.g. an antibody fragment.  
 CC (Updated on 09-JAN-2003 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 267 AA;  
 Query Match 100.0%; Score 1020; DB 14; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-100;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLLTGADVNRHEIPVLPNRYGLPINORFTLV 60  
 DB 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLLTGADVNRHEIPVLPNRYGLPINORFTLV 60  
 QY 61 ELSNHAELSVTLADVTNAYVVGIRAGNSAYFFHPDQEDAEATHLFTDVONRYTFARF 120  
 DB 61 ELSNHAELSVTLADVTNAYVVGIRAGNSAYFFHPDQEDAEATHLFTDVONRYTFARF 120  
 QY 121 GNYDRLEQLAGNLRNIEELGNGPPEEASISALYYSTGQTLPLARSFIICMISEAR 180  
 DB 121 GNYDRLEQLAGNLRNIEELGNGPPEEASISALYYSTGQTLPLARSFIICMISEAR 180  
 QY 181 FOYIEGEMRTIRIRYRRS 198  
 DB 181 FOYIEGEMRTIRIRYRRS 198  
 RESULT 2  
 ID AAR63902 standard; protein; 267 AA.  
 AC AAR63902;  
 XX  
 AC 25-MAR-2003 (updated)  
 DT 27-JUL-1995 (first entry)  
 XX  
 DE Ricin A-chain (RTA).  
 XX  
 KW Ricin A chain: RTA: ribosome-inactivating proteins; RIPS;  
 KW cytotoxic therapeutic agents; autoimmune disease; cancer;  
 KW graft-versus-host disease.  
 XX  
 OS Ricinus communis.  
 OS  
 PN W09426910-A1.  
 PD 24-NOV-1994.  
 XX  
 PF 12-MAY-1994; 94WO-US05348.  
 XX  
 PR 12-MAY-1993; 93US-0064691.  
 XX  
 PA (XOMA ) XOMA CORP.  
 XX  
 PI Better MD, Carroll SS, Studnicka GM, Carroll SF;  
 DR WPI; 1995-006804/O1.  
 XX  
 PT Polynucleotide(s) encoding type I ribosome-inactivating proteins

PT - which are suitable for use as components of cytotoxic  
 PT therapeutic agents.  
 XX  
 PS Example 3; Fig 1; 221pp; English.  
 XX  
 PS AAR63902 is the ricin A chain gene product, it is analogous to the  
 CC ribosome-inactivating proteins (RIPs) described in AAR63903-R63911.  
 CC RIPs are the key components of cytotoxic therapeutic agents (CTAs),  
 CC which include gene fusion products and immunconjugates. CTAs may  
 CC be used to selectively eliminate any cell type to which a RIP  
 CC component is targeted, by the specific binding capacity of the  
 CC second component of the agent. They can be used in the treatment  
 CC of diseases where the elimination of a particular cell type is  
 CC desired, such as autoimmune disease, cancer and graft-versus-host  
 CC disease.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 267 AA;  
 Query Match 100.0%; Score 1020; DB 16; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-100;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLLTGADVNRHEIPVLPNRYGLPINORFTLV 60  
 DB 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLLTGADVNRHEIPVLPNRYGLPINORFTLV 60  
 QY 61 ELSNHAELSVTLADVTNAYVVGIRAGNSAYFFHPDQEDAEATHLFTDVONRYTFARF 120  
 DB 61 ELSNHAELSVTLADVTNAYVVGIRAGNSAYFFHPDQEDAEATHLFTDVONRYTFARF 120  
 QY 121 GNYDRLEQLAGNLRNIEELGNGPPEEASISALYYSTGQTLPLARSFIICMISEAR 180  
 DB 121 GNYDRLEQLAGNLRNIEELGNGPPEEASISALYYSTGQTLPLARSFIICMISEAR 180  
 QY 181 FOYIEGEMRTIRIRYRRS 198  
 DB 181 FOYIEGEMRTIRIRYRRS 198  
 RESULT 3  
 ID AAW25136 standard; protein; 290 AA.  
 AC AAW25136;  
 XX  
 AC 25-MAR-2003 (updated)  
 DT 02-DEC-1997 (first entry)  
 XX  
 DE Ricin A-chain ribosome inhibitory protein inactive precursor.  
 XX  
 KW Matze; proRIP: ribosome inactivating protein; alpha; beta subunit;  
 KW internal linker; Barley Translation Inhibitor; Trichosanthin;  
 KW Ricin A-chain; Abirin-A-A-chain; Saporin; SIT-1; Luffin A; MAP;  
 KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;  
 KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;  
 KW post-translational modification; cancer; neoplasia; HIV; AIDS;  
 KW human immunodeficiency virus; acquired immune deficiency syndrome.  
 XX  
 OS Synthetic.  
 OS  
 PN US5646026-A.  
 PD 08-JUL-1997.  
 XX  
 PF 07-JUN-1995; 95US-0485286.  
 XX  
 PR 09-DEC-1992; 92US-0987927.  
 PR 11-JUN-1990; 90US-0535636.  
 PR 26-JAN-1995; 95US-0378761.  
 PR 07-JUN-1995; 95US-0485286.  
 XX  
 PA (DOWC ) DOWELANCO.

XX Hey TD, Morgan AER, Walsh TA;  
 PI WPI; 1997-362934/33.  
 DR  
 XX  
 PT DNA encoding pro-ribosome inactivating proteins - inactive  
 PT precursors of ribosome inactivating proteins; can be expressed in  
 PT eukaryotic cells without causing cell death  
 XX  
 PS Claim 4; Column 91-94; 186pp; English.  
 CC  
 XX AAM25136 represents a RlcA A-chain ribosome inhibitory protein (RIP)  
 CC which was engineered to contain a selectively removable internal peptide  
 CC linker sequence separating the alpha and beta units of the RIP. When  
 CC separated the two units regain activity and are capable of inactivating  
 CC eukaryotic ribosomes and hence preventing protein production. Many  
 CC different RIPs may be produced with an internal linker including  
 CC maize RIP, Trichosanthin, RlcA A-chain, Abirin-A A-chain and  
 CC Saporin. The RIPs can be used in the construction of therapeutic  
 CC toxins targeted to specific cells such as tumour cells via the  
 CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.  
 CC A further use is in HIV therapy (see US4869903). There is interest  
 CC in expressing RIP recombinantly in host eukaryotic cells, because of  
 CC the capacity to provide correct post-translational processing. However,  
 CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting  
 CC in cell death. Since the inactive RIP proteins are not cytotoxic to  
 CC eukaryotic cells, they can be recombinantly expressed in such cells and  
 CC then converted to active RIP proteins.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SO Sequence 290 AA;  
 Query Match 100.0%; Score 1020; DB 18; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-100;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IFPKQYPIINFTTAGATVOSTYTNFIRAVRGRLTTGADVREHIEPVLNPRVGLPINORFTLV 60  
 DB 25 IFPKQYPIINFTTAGATVOSTYTNFIRAVRGRLTTGADVREHIEPVLNPRVGLPINORFTLV 84  
 QY 61 ELSNHAELSVTLALDVTNNAVYVGRAGNSAYFFPHNDEDAEATITLFTDVONRYTFAG 120  
 DB 85 ELSNHAELSVTLALDVTNNAVYVGRAGNSAYFFPHNDEDAEATITLFTDVONRYTFAG 144  
 QY 121 GNYDRLEQLAGNLRENIEIGNGPLEEASALYYSTGCTQLPTLARSFTICIMISEAR 180  
 DB 145 GNYDRLEQLAGNLRENIEIGNGPLEEASALYYSTGCTQLPTLARSFTICIMISEAR 204  
 QY 181 FOYIEGEMRTIRIRNRRS 198  
 DB 205 FOYIEGEMRTIRIRNRRS 222  
 RESULT 4  
 AAM21699  
 ID AAM21699 standard; Protein; 290 AA.  
 AC AAM21699;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 26-SEP-1997 (first entry)  
 XX  
 DE RlcA A-chain RIP.  
 XX  
 KM pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;  
 KM inactivator; eukaryotic ribosome; alpha fragment; beta fragment;  
 KM inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;  
 KM rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.  
 XX  
 OS RlcA A-chain communis.  
 XX  
 XX  
 FH Key  
 FT Region 152..162

PT /note= "Position of possible insertion of internal  
 PT peptide linker sequence"  
 XX  
 PN US5635384-A.  
 XX  
 PD 03-JUN-1997.  
 XX  
 PF 26-JAN-1995; 9505-0378761.  
 XX  
 PR 09-DEC-1992; 92US-0987927.  
 PR 11-JUN-1990; 90US-0535636.  
 XX 26-JAN-1995; 9505-0378761.  
 XX  
 PA (DOMC ) DOWELANCO.  
 XX  
 PI Hey TD, Morgan AER, Walsh TA;  
 XX WPI; 1997-309831/28.  
 DR  
 XX  
 XX Inactive precursor of maize ribosome-inactivating protein - also  
 PT chimeric ribosome-inactivating protein precursors containing  
 PT internal linker sequences  
 XX  
 PS Claim 2; Column 91-94; 121pp; English.  
 XX  
 CC The sequences given in AAM21698-710 represent Ribosome Inactivating  
 CC Proteins (RIP's), which may be used in the construction of the  
 CC proRIP of the invention. The proRIP has a selectively removable,  
 CC internal peptide linker. The precursor sequence is incapable of  
 CC inactivating eukaryotic ribosomes, but can be converted by removal  
 CC of the linker into a protein having alpha and beta fragments and being  
 CC capable of inactivating eukaryotic ribosomes. RIPs are potent  
 CC inhibitors of eukaryotic protein synthesis. They possess a highly  
 CC specific N-glycosidase activity which cleaves the glycosidic bond of  
 CC adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit  
 CC cellular proliferation of cells, e.g. cancer cells and HIV-infected T  
 CC cells. The inactive proRIP proteins make it possible to provide protein  
 CC synthesis inhibitors with uses in practical and improved ways not before  
 CC possible. The RIP can be used to make cytotoxic conjugates.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SO Sequence 290 AA;  
 Query Match 100.0%; Score 1020; DB 18; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-100;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IFPKQYPIINFTTAGATVOSTYTNFIRAVRGRLTTGADVREHIEPVLNPRVGLPINORFTLV 60  
 DB 25 IFPKQYPIINFTTAGATVOSTYTNFIRAVRGRLTTGADVREHIEPVLNPRVGLPINORFTLV 84  
 QY 61 ELSNHAELSVTLALDVTNNAVYVGRAGNSAYFFPHNDEDAEATITLFTDVONRYTFAG 120  
 DB 85 ELSNHAELSVTLALDVTNNAVYVGRAGNSAYFFPHNDEDAEATITLFTDVONRYTFAG 144  
 QY 121 GNYDRLEQLAGNLRENIEIGNGPLEEASALYYSTGCTQLPTLARSFTICIMISEAR 180  
 DB 145 GNYDRLEQLAGNLRENIEIGNGPLEEASALYYSTGCTQLPTLARSFTICIMISEAR 204  
 QY 181 FOYIEGEMRTIRIRNRRS 198  
 DB 205 FOYIEGEMRTIRIRNRRS 222  
 RESULT 5  
 AAP70097  
 ID AAP70097 standard; protein; 332 AA.  
 AC AAP70097;  
 XX  
 DT 09-APR-1991 (first entry)  
 XX  
 DE RlcA A.

```

XX  Ricin A; Met-aminopeptidase.
KW  Escherichia coli.
OS  EP219237-A.
PN  22-APR-1987.
PD  19-SEP-1986; 86EP-0307242.
XX  06-MAY-1986; 86US-0860330.
PR  20-SEP-1985; 85US-0778414.
XX  (CETU ) CETUS CORP.
PA  Benbasat A, Bauer KA, Chang S, Chang SY;
PI  WPI: 1987-110172/16.
DR  N-PSDB; AAN70152.
XX  N-terminal methionine free proteins prodn. - by using host
PT  transformed with vector to express a methionine-amino-peptidase
PS  Disclosure; Fig. 4; 20pp; English.
XX  Ricin A may be produced in a form which lacks an N-terminal Met
CC  using Met-aminopeptidase from E.coli.
XX  Sequence 332 AA;
SQ  Query Match 100.0%; Score 1020; DB 8; Length 332;
    Best Local Similarity 100.0%; Pred. No. 6.5e-100;
    Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 IFFKQYPIINFTTAGATVOSTYTNFIRAVRGRLTTGADVHHEIPVLPNRGGLPINQRETLV 60
    |||||||
DB  36 IFFKQYPIINFTTAGATVOSTYTNFIRAVRGRLTTGADVHHEIPVLPNRGGLPINQRETLV 95
QY  61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPNODDAEATITLFTDVONRYTFEAFG 120
    |||||||
DB  96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPNODDAEATITLFTDVONRYTFEAFG 155
QY  121 GNYDRLEQLAGNLRENIELGNGPLLEAISAALYYSTGGTQLPTLARSFTICIQMISEAAR 180
    |||||||
DB  156 GNYDRLEQLAGNLRENIELGNGPLLEAISAALYYSTGGTQLPTLARSFTICIQMISEAAR 215
QY  181 FOYIEGEMRTIRIRYNRRS 198
    |||||||
DB  216 FOYIEGEMRTIRIRYNRRS 233

RESULT 6
AAP70838
ID  AAP70838 standard; protein; 332 AA.
XX
AC  AAP70838;
XX
DT  25-MAR-2003 (updated)
DT  18-FEB-1991 (first entry)
XX
DE  Sequence of Ricinus communis castor beans ricin toxin (RT or ricin)
DE  A protein encoded by PRA123.
XX
KW  Lectin; toxin protein; cytotoxic; cytostatic; castor bean;
KW  plant toxin.
XX
OS  Ricinus communis.
XX
XX  Key Location/Qualifiers
FH  1..32
FT  /note="Leader"
FT  Region 33..302
FT  Region

```

```

FT  /note="A-chain"
FT  Region 315..332
FT  /note="B-chain"
XX  EP237676-A.
XX  23-SEP-1987.
XX  13-NOV-1986; 86EP-0308877.
XX  07-MAR-1986; 86US-0837583.
XX  (CETU ) CETUS CORP.
XX  (CHIR ) CHIRON CORP.
XX  Platak M;
XX  WPI: 1987-265177/38.
DR  N-PSDB; AAN70519.
XX  New non-glycosylated ricin precursor and toxin etc. - are prep'd.
PT  by recombinant DNA procedures with specific isolation steps for
PS  purer and soluble prods.
XX  Disclosure; Fig 1; 112pp; English.
XX  The full-length sequences encoding ricin A (AAN70520), ricin D
CC  (AAN70525), putative ricin E (AAN70526) and RCA (AAN70524) in precursor
CC  form were obt'd. using messenger RNA to obtain a cDNA library, and
CC  then probing the library to retrieve the desired cDNA inserts. The
CC  library was probed using the 35-mer given in AAN70514. Figure 4 (see
CC  AAN70520, AAN70521, AAN70522) shows the nucleotide sequences of three
CC  plasmids contg. cDNA inserts obt'd. by probing a cDNA library for
CC  sequences encoding ricin B using the probe in AAN70517. The cDNA
CC  inserts can be placed into expression vectors. Site-directed
CC  mutagenesis may be used to place an ATG start codon and a HindIII
CC  site at the beginning of the mature protein (see AAN70518). The
CC  coding sequences of the inserts can be ligated into expression
CC  vectors contg. the phoA promoter-operator and leader sequence
CC  (AAN70523) and suitable retroregulators.
XX  (updated on 25-MAR-2003 to correct PA field.)
XX  Sequence 332 AA;
SQ  Query Match 100.0%; Score 1020; DB 8; Length 332;
    Best Local Similarity 100.0%; Pred. No. 6.5e-100;
    Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 IFFKQYPIINFTTAGATVOSTYTNFIRAVRGRLTTGADVHHEIPVLPNRGGLPINQRETLV 60
    |||||||
DB  36 IFFKQYPIINFTTAGATVOSTYTNFIRAVRGRLTTGADVHHEIPVLPNRGGLPINQRETLV 95
QY  61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPNODDAEATITLFTDVONRYTFEAFG 120
    |||||||
DB  96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPNODDAEATITLFTDVONRYTFEAFG 155
QY  121 GNYDRLEQLAGNLRENIELGNGPLLEAISAALYYSTGGTQLPTLARSFTICIQMISEAAR 180
    |||||||
DB  156 GNYDRLEQLAGNLRENIELGNGPLLEAISAALYYSTGGTQLPTLARSFTICIQMISEAAR 215
QY  181 FOYIEGEMRTIRIRYNRRS 198
    |||||||
DB  216 FOYIEGEMRTIRIRYNRRS 233

RESULT 7
AAP95639
ID  AAP95639 standard; protein; 332 AA.
XX
AC  AAP95639;
XX
DT  25-MAR-2003 (updated)
DT  31-OCT-2002 (updated)

```



DT	13-AUG-1990	(first entry)
DE	Ricin A encoded by insert from plasmid pRA123.	
XX		
KW	Plasmid pRA123; ricin-A; ricin-B; cytotoxicity.	
XX		
OS	Ricinus communis.	
XX		
XX	Synthetic.	
EH	Key	
EH	Location/Qualifiers	
FT	Peptide	
FT	1..35	
FT	/label= leader sequence	
FT	36..302	
FT	/label=A-chain	
FT	303..314	
FT	/label=linker	
FT	315..332	
FT	/label=B-chain	
XX		
XX	EP335476-A.	
XX		
PD	04-OCT-1989.	
XX		
XX	19-JAN-1989; 89EP-0201162.	
XX		
PR	08-FEB-1984; 84US-0578115.	
PR	08-FEB-1984; 84US-0578121.	
PR	09-FEB-1984; 84US-0578122.	
PR	07-SEP-1984; 84US-0648759.	
PR	20-SEP-1984; 84US-0653515.	
XX		
PA	(CETU ) CETUS CORPORATION.	
PI		
PI	Gelfand D, Lawyer FC, Horn G, Greenfield L, Nitecki D, Kaplan D;	
PI	Platak MJ;	
XX		
DR	WPI; 1989-286959/40.	
DR	N-PSDB; AAN91281.	
XX		
PT	Recombinant vectors expressing ricin chains or diphtheria toxin -used for	
PT	prodn. Of new immunotoxin conjugates with monoclonal antibodies, having	
XX	high cell specificity and good extracellular stability.	
XX		
XX	Disclosure; Fig 14; 54pp; English.	
XX		
CC	Plasmid RA123 (ATCC No. 39799) carries the entire coding sequence for	
CC	ricin A, as well as codons for 12 AAs joining the A to the B chain.	
CC	Following modification for ease of manipulation the plasmid was used to	
CC	construct expression vectors which express the conjugates in	
CC	host cells.	
CC	(Updated on 31-OCT-2002 to add missing OS field.)	
CC	(Updated on 25-MAR-2003 to correct PR field.)	
CC	(Updated on 25-MAR-2003 to correct PR field.)	
CC	(Updated on 25-MAR-2003 to correct PI field.)	
XX		
XX	Sequence 332 AA:	
XX		
Query Match	100.0%; Score 1020; DB 10; Length 332;	
Best Local Similarity	100.0%; Pred. No. 6.5e-100;	
Matches 198; Conservative	0; Mismatches 0; Indels 0; Gaps 0	
1	IFPKQYPIINFETAGATVQSYTNFTRAVGRITTGADVYHEITPVLPNRYGLPIINOFILY 60	
36	IFPKQYPIINFETAGATVQSYTNFTRAVGRITTGADVYHEITPVLPNRYGLPIINOFILY 95	
61	ELSNRAELSVTLALDYTNAYVGYGRAGNSAIFPHDPNDEDAEALITHLFTDVONRTYFAG 120	
96	ELSNRAELSVTLALDYTNAYVGYGRAGNSAIFPHDPNDEDAEALITHLFTDVONRTYFAG 155	
121	GVNDEDELAGULRENIENGNGPLLEAISAALYYSGGQQLPTLARSFIIICOMISEAR 180	
156	GVNDEDELAGULRENIENGNGPLLEAISAALYYSGGQQLPTLARSFIIICOMISEAR 215	

QY	181	FOYIEGEMRTIRYRRRS	198	
DB	216	FOYIEGEMRTIRYRRRS	233	
RESULT 8				
ID	AA070827			
XX	AA070827	standard; Protein; 554	AA.	
AC	AA070827;			
XX				
DT	25-MAR-2003	(updated)		
DI	31-AUG-1995	(first entry)		
XX				
DE	Anti-cataract immunotoxin.			
XX				
KW	Immunotoxin; heavy chain; light chain; variable region; antibody;			
KW	ricin-A; cytostatic; cataract; lens opacification; epithelial cell;'			
KW	PHB19; 4197X; monoclonal antibody; Mab.			
XX				
OS	Synthetic.			
XX				
FH	Key	Location/Qualifiers		
FT	Peptide	1..27		
FT		/label= Sig_peptide		
FT		/note= "phoA signal sequence"		
FT	Domain	28..145		
FT		/label= HEAVY		
FT	Peptide	/note= "Mab 4197X heavy chain"		
FT		148..166		
FT	Domain	/label= LINKER		
FT		169..274		
FT		/label= LIGHT		
FT		/note= "Mab 419X light chain"		
FT	Domain	276..544		
FT		/label= RICIN-A		
FT	Peptide	549..554		
FT		/label= TAG		
FT		/note= "hexa-histidine tail"		
XX				
PN	W09503828-A1.			
XX				
PD	09-FEB-1995.			
XX				
PF	15-JUL-1994;	94WO-US07919.		
XX				
PR	02-AUG-1993;	93US-0101329.		
XX				
PA	(HOUS-) HOUSTON BIOTECHNOLOGY INC.			
XX				
PI	Gould RM, Kelleher PJ, Wallace TL, Wood MS;			
XX				
DR	WPI; 1995-082036/11.			
DR	N-PSDB; AA085386.			
XX				
XX	New single chain immuno:toxin - binds specifically to epithelial			
PT	cells, for inhibiting development of sec. cataracts after			
PT	extra:capsular cataract extraction.			
XX				
PS	Disclosure: Fig.4; 68pp; English.			
XX				
CC	The immunotoxin given in AAR70827 comprises the heavy and light chain			
CC	variable regions of anti-lens epithelium IgG3 Mab 4197X linked to			
CC	ricin-A and a hexa-histidine tag. The DNA construct encoding the			
CC	immunotoxin was expressed from PHB19 in E. coli.			
CC	(Updated on 25-MAR-2003 to correct PN field.)			
XX				
SO	Sequence	554	AA;	
Query Match 100.0%; Score 1020; DB 16; Length 554;				
Best Local Similarity 100.0%; Pred. No. 1,3e-19;				
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				

```

OY 1 IFPKOYPIINFTAGATVOSTYNTFIRAVRGLTTGADVRHETPLVPLNVRGLPINORFTLV 60
DB 278 IFPKOYPIINFTAGATVOSTYNTFIRAVRGLTTGADVRHETPLVPLNVRGLPINORFTLV 337
OY 61 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAELTHLFTDVQNRFTAFG 120
DB 338 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAELTHLFTDVQNRFTAFG 397
OY 121 GNYDRLEQLAGNLRNENIELGNPLEEALISALYYSTGQTOLPTLARSFTICIQMISEAR 180
DB 398 GNYDRLEQLAGNLRNENIELGNPLEEALISALYYSTGQTOLPTLARSFTICIQMISEAR 457
OY 181 FOYIEGEMRTIRYNNRS 198
DB 458 FOYIEGEMRTIRYNNRS 475

RESULT 9
AAP90079
ID AAP90079 standard; protein; 562 AA.
AC
XX AAP90079;
AC
XX 25-MAR-2003 (updated)
DT 01-NOV-1989 (first entry)
XX
DE Ricin D.
XX
KW Ricin D; Ricinus communis; castor beans; Zanibaricensis variety;
KW modified; lectin binding removed; reduced cell binding
XX
OS Ricinus communis (castor beans).
XX
PN WO8904839-A.
XX
PD 01-JUN-1989.
XX
PF 23-NOV-1988; 88WO-US04238.
XX
PR 24-NOV-1987; 87US-0124735.
XX
PA (GEMV) GENETICS INST INC.
XX
PI Brown EL, Jones S;
XX
DR WPI; 1989-178366/24.
DR N-PSDB; AAN90068.
XX
PT Modified ricin molecules and toxin conjugates
PT - in which the lectin binding function of the B chain
PT is removed or diminished to reduce cell binding.
XX
PS Disclosure; fig 1; 51pp; English.
XX
XX Ricin D (see corresp. AAN90068) found on the EcoRI-HindIII fragment
CC of DNA from Ricinu communis, Zanibaricensis variety. Patent
CC discloses many modifications of ricin in which the lectin binding
CC function of the B chain is diminished or removed, and conjugation
CC to toxins to eliminate cell binding.
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 562 AA;

Query Match 100.0%; Score 1020; DB 10; Length 562;
Best Local Similarity 100.0%; Pred. No. 1,3e-99;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 96 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAELTHLFTDVQNRFTAFG 155
OY 121 GNYDRLEQLAGNLRNENIELGNPLEEALISALYYSTGQTOLPTLARSFTICIQMISEAR 180
DB 156 GNYDRLEQLAGNLRNENIELGNPLEEALISALYYSTGQTOLPTLARSFTICIQMISEAR 215
OY 181 FOYIEGEMRTIRYNNRS 198
DB 216 FOYIEGEMRTIRYNNRS 233

RESULT 10
AAG78304
ID AAG78304 standard; protein; 565 AA.
AC
XX AAG78304;
AC
XX 27-NOV-2001 (first entry)
DT
XX
DE Modified castor bean preprotricin (SEQ ID 10).
XX
XX Castor bean plant; preprotricin; ricin; A chain; B chain;
KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;
KW retroviral infection; anti-HIV; virucide; viral protease.
XX
OS Chimeric - Ricinus communis
OS Chimeric - Human immunodeficiency virus type 2.
XX
FH Key
FH Peptide
FT 1..24
FT /label= Signal_peptide
FT 25..565
FT /label= Protrich
FT /note= "Protrich consists of the ricin A chain, a linker
FT peptide, and the ricin B chain. Protrich is
FT proteolytically cleaved between the A chain and
FT the linker to yield mature ricin"
FT
FT Protein
FT 25..291
FT /label= Ricin_A_chain
FT /note= "N-glycosidase"
FT 292..303
FT /label= Linker_peptide
FT 296..297
FT /label= HIV_protease_cleavage_site
FT 304..565
FT /label= Ricin_B_chain
FT /note= "Galactose/N-acetylglactosamine-binding lectin"
XX
XX WO200160393-A1.
XX
XX 23-AUG-2001.
XX
XX 15-FEB-2001; 2001WO-US05282.
XX
XX 16-FEB-2000; 2000US-0182759.
XX
XX (BECH-) BECHTEL BWXT IDAHO LLC.
XX
XX Keener WK, Ward TE;
XX
XX WPI; 2001-581908/65.
XX
XX N-PSDB; AAI64145.
XX
XX Novel composition comprising toxin e.g., ricin based antiviral compound
PT useful for treating viral infections such as human immunodeficiency
PT virus infection.
XX
XX Example 1; Page 59-63; 66pp; English.
XX
XX The sequence relates to the amino acid sequence of a modified preprotricin
XX protein encoded by AAI64145. The invention relates to a novel toxin
XX (e.g. ricin) based antiviral agent which is toxic to virus-infected
CC

```

XX	PI	Plataek M;
XX	DR	WPI; 1987-265177/38.
XX	DR	N-PSDB; AAN70526.
XX	PT	New non-glycosylated ricin precursor and toxin etc. - are prepd.
XX	PT	by recombinant DNA procedures with specific isolation steps for
XX	PT	purifier and soluble prods.
XX	PS	Disclosure; Fig 14(1-2); 112pp; English.
XX	CC	The full length sequences encoding ricin A (AAN70520), ricin D
XX	CC	(AAN70525) putative ricin E (AAN70526) and RCA (AAN70524) in precursor
XX	CC	form were obtained, using the messenger RNA to obtain a cDNA library, and
XX	CC	then probing the library to retrieve the desired cDNA inserts. The
XX	CC	library was probed using the 35-mer given in AAN70514, Figure 4 (see
XX	CC	AAN70520, AAN70521, AAN70522), shows the nucleotide sequences of three
XX	CC	plasmids containing cDNA inserts obtained by probing a cDNA library
XX	CC	for sequences encoding ricin B using the probe in AAN70517. The cDNA
XX	CC	inserts can be placed into expression vectors. Site-directed
XX	CC	mutagenesis may be used to place an ATG start codon and a HindIII
XX	CC	site at the beginning of the mature protein. (see AAN70518). The
XX	CC	coding sequences of the inserts can be ligated into expression
XX	CC	vectors containing the PhdA promoter-operator and leader sequence
XX	CC	(AAN70523) and suitable retroregulators.
XX	CC	(Updated on 25-MAR-2003 to correct PA field.)
XX	SQ	Sequence 576 AA;
XX	Query Match	100.0%; Score 1020; DB 8; Length 576;
XX	Best Local Similarity	100.0%; Pred. No. 1.4e-99;
XX	Matches 198; Conservative	0; Mismatches 0; Indels 0; Gaps 0
XX	OY	1 IFPKOYPIINFTTAGATVOSYTNFTRAVRGRLLTGADVRHEIPVLPRNGLPINORFTLV 60
XX	Db	36 IFPKOYPIINFTTAGATVOSYTNFTRAVRGRLLTGADVRHEIPVLPRNGLPINORFTLV 95
XX	OY	61 ELSNHELSTLTALDVNTNAYVYVGRAGNSAFFPHPDNOEAEATHTLFTVONRRTFARG 120
XX	Db	96 ELSNHELSTLTALDVNTNAYVYVGRAGNSAFFPHPDNOEAEATHTLFTVONRRTFARG 155
XX	OY	121 GNYDRLEOLAGNLRNIEIENGPLEEASALYYSTGGTOLPTLARSFTICIONISEAR 180
XX	Db	156 GNYDRLEOLAGNLRNIEIENGPLEEASALYYSTGGTOLPTLARSFTICIONISEAR 215
XX	OY	181 FOYIEGDMRTIRIRYNRS 198
XX	Db	216 FOYIEGDMRTIRIRYNRS 233
XX	RESULT 12	
XX	AAM25787	
XX	ID	AAM25787 standard; Protein: 576 AA.
XX	XX	AAM25787;
XX	AC	
XX	XX	25-MAR-2003 (updated)
XX	DT	27-MAR-1998 (first entry)
XX	XX	
XX	XX	Castorbean ricin.
XX	XX	
XX	XX	Ricin; cytotoxin; hybrid protein; cell delivery;
XX	KM	cell binding ligand; translocation domain; diphtheria toxin B';
XX	KM	interleukin-2; T-cell lymphoma; organ rejection; therapy.
XX	OS	Ricinus communis.
XX	XX	
XX	Key	Location/Qualifiers
XX	FT	1..35
XX	FT	/label= Sig_peptide
XX	FT	36..302
XX	FT	/label= A-domain
XX	FT	Protein



Query Match 100.0%; Score 1020; DB 20; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-99;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFKQYPIINFTAGATVOSTNFTIRAVGRGLTTGADVREHEIPVLPNRYGLPIINORFLLV 60  
 |||||  
 DB 36 IFFKQYPIINFTAGATVOSTNFTIRAVGRGLTTGADVREHEIPVLPNRYGLPIINORFLLV 95  
 |||||  
 QY 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEITLHFTDVONRYTFAFG 120  
 |||||  
 DB 96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEITLHFTDVONRYTFAFG 155  
 |||||  
 QY 121 GNYDRLEQLAGNRENIELGNGPLEAISALYYSTGQTOLPTLARSFIICQMISEAR 180  
 |||||  
 DB 156 GNYDRLEQLAGNRENIELGNGPLEAISALYYSTGQTOLPTLARSFIICQMISEAR 215  
 |||||  
 QY 181 FOYIEGEMRTIRIRYRNR 198  
 |||||  
 DB 216 FOYIEGEMRTIRIRYRNR 233

RESULT 14  
 AAY78592  
 ID AAY78592 standard; Protein: 576 AA.

AC AAY78592;  
 XX  
 DT 05-MAY-2000 (first entry)

XX Ricinus communis ricin protein sequence.

KW Ricin; toxin; hybrid protein; translocation domain; cell destruction;  
 KW cell binding domain; genetic deficiency disease; cell targeting; cancer;  
 KW adipocyte; enzyme delivery; anti-viral; HIV.

XX Ricinus communis.

XX US6022950-A.

PD 08-FEB-2000.

PE 07-JUN-1995; 95US-0479510.

PR 07-JUN-1984; 84US-0618199.

PR 27-JUN-1991; 91US-0722484.

PR 25-APR-1985; 85US-0726808.

PR 07-JUN-1985; 85US-0742554.

PR 22-DEC-1989; 89US-0456095.

PR 14-JUN-1990; 90US-0538276.

PR 04-AUG-1993; 93US-0102387.

PA (SERA-) SERAGEN INC.

PI Murphy JR.

XX WPI: 2000-160390/14.

XX N-PSDB: AAZ90019.

XX New two-part hybrid protein comprising a translocation domain and a

XX cell-binding domain, for treating genetic deficiency diseases, cancer

XX and HIV infections -

XX Example 4; Fig 11; 32pp; English.

XX This sequence represents the Ricinus communis ricin protein sequence. The

XX toxin can be included in the hybrid protein of the invention and used to

XX destroy or modify the cell that the hybrid protein is targeted to. The

XX hybrid protein comprises a first part which is a portion of the binding

XX domain of a cell-binding ligand, effective to cause the hybrid molecule

XX to bind to a cell of an animal. The second part comprises a portion of a

XX translocation domain of a naturally occurring protein (e.g. the

XX translocation domain of diphtheria toxin) the second part translocates

XX the third part across the cytoplasmic membrane and into the cytosol of

CC the cell. The third part comprises a chemical entity to be introduced  
 CC into the cell, where each of the first and third part is non-native with  
 CC respect to naturally occurring protein, and the covalent bond attaching  
 CC the second and third part is cleavable. The toxin represented by the  
 CC present sequence can form part of the third portion of the hybrid  
 CC protein. The cell binding domain binds to a specific cell and the  
 CC translocation domain transfers the hybrid molecule across the cell  
 CC membrane into the cytosol. The third part of the protein, linked to the  
 CC translocation domain through a cleavable bond, can then carry out its  
 CC function. The hybrid molecules are useful for treating genetic deficiency  
 CC diseases by delivering to affected cells an enzyme supplying the missing  
 CC function, to supplement cellular levels of a particular enzyme or a  
 CC scarce precursor or cofactor, to direct toxins or other poisons to  
 CC destroy particular cells (such as adipocytes, cancer cells, or  
 CC virus-infected cells), and to counteract viral infections such as HIV by  
 CC introducing into appropriate cells antibodies to viral proteins.

XX Sequence 576 AA;

Query Match 100.0%; Score 1020; DB 21; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-99;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFKQYPIINFTAGATVOSTNFTIRAVGRGLTTGADVREHEIPVLPNRYGLPIINORFLLV 60  
 |||||

DB 36 IFFKQYPIINFTAGATVOSTNFTIRAVGRGLTTGADVREHEIPVLPNRYGLPIINORFLLV 95  
 |||||

QY 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEITLHFTDVONRYTFAFG 120  
 |||||

DB 96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEITLHFTDVONRYTFAFG 155  
 |||||

QY 121 GNYDRLEQLAGNRENIELGNGPLEAISALYYSTGQTOLPTLARSFIICQMISEAR 180  
 |||||

DB 156 GNYDRLEQLAGNRENIELGNGPLEAISALYYSTGQTOLPTLARSFIICQMISEAR 215  
 |||||

QY 181 FOYIEGEMRTIRIRYRNR 198  
 |||||

DB 216 FOYIEGEMRTIRIRYRNR 233

RESULT 15  
 AAG78301  
 ID AAG78301 standard; Protein: 576 AA.

AC AAG78301;

XX 15-NOV-2001 (first entry)

XX Castor bean preproprotein (SEQ ID 2).

XX Castor bean plant; preproprotein; ricin; A chain; B chain;

XX human immunodeficiency virus infection; HIV; toxin; antiviral agent;

XX retroviral infection; anti-HIV; virucide activity; viral protease.

XX Ricinus communis.

OS Key Location/Qualifiers

XX Peptide 1..35

XX Protein /label= Signal peptide

XX Peptide /label= Ricin\_A\_chain

XX Protein /note= "N-glycosidase"

XX Peptide /label= Linker\_peptide

XX Protein /note= "Cleaved during activation of ricin"

XX Protein /label= Ricin\_B\_chain

XX Protein /note= "Galactose/N-acetylgalactosamine-binding lectin"

XX WO200160393-A1.

XX 23-AUG-2001.

PF 15-FEB-2001; 2001WO-US05282.

PR 16-FEB-2000; 2000US-0182759.

PA (BECH-) BECHTEL BWXT IDAHO LLC.

PI Keener WK, Ward TE;

DR WPI; 2001-581908/65.

XX  
X

X  
X  
X

PT	Novel composition comprising toxin e.g., ricin based antiviral compound
PT	useful for treating viral infections such as human immunodeficiency
PT	virus infection.

PS Disclosure; Page 50-54; 66pp; English.

CC The sequence relates to preporcricin protein encoded by the DNA sequence  
CC given in MA164138. The invention relates to a novel toxin (e.g., ricin)  
CC based antiviral agent which is toxic to virus-infected cells, but  
CC non-toxic to uninfected cells. The invention has anti-HIV and virucide  
CC activities. Its mechanism of action is through inactivation of cellular  
CC ribosomes and enhancement of binding of the antiviral agent to galactose  
CC residues on cell surfaces, and its cellular internalisation. The  
CC invention is useful for treating human immunodeficiency virus infection  
CC and other viral infections, especially retroviral infections. The  
CC antiviral agent is activated in viral particles or early-stage infected  
CC cells, killing the cells upon infection and effectively preventing the  
CC integration of the viral genome into the host genome thereby preventing  
CC the latency/rebound problem. The agent enters all HIV susceptible cells,  
CC and not just cells known to act as host cells for the virus. The  
CC antiviral agent remains inert in a cell until degraded in it, unless the  
CC cell is infected with the virus, where the viral protease activates it.

**SQ Sequence 576 AA;**

Query Match

100.0%; Score 1020; DB 22; Length 576;

Best Local Similarity 100.0%; Pred. NO. 1.4e-99;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINQRFILV 60

Db 36 IFPKQYPIINF<sup>T</sup>TAGATVQSYTNFIRAVRGRLTTGADV<sup>R</sup>HEIPVLPN<sup>R</sup>VGLPINQ<sup>R</sup>FI<sup>L</sup>V 95

61 ELSNHAELSVTLALDVTNAYVVG YRAGNSAYFFHPDNQEDAEAITHLFTDVQNRRTFAFG 120

Db 96 ELSNHAELSVTLALDVTNAYVVG YRAGNSAYFFHPDNQEDAEAITHLFTDVQNRRTFAFG 155

121 GNYDRLEQLAGNLRNIEICNGPLEEASALYYSTGGTQLPTLARSFIICIMISEAR 180

D6 156 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYSTGGTQLPTLARSFIIQIMISEAR 213

181 FOYIEGEMTRIRYNRRS 198  
QY

Db 216 FQY1EGEMRTIRIKYNKRS 233

Search completed: September 16, 2003, 11:45:17  
Job time : 33.9711 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:30:32 ; Search time 6.94128 Seconds  
(without alignments)  
1341.437 Million cell updates/sec

Title: US-10-083-336A-3

Perfect score: 1020  
Sequence: 1 IFPKQYPIINFTTAGATVQS.....ARFQYIEGEMRTIRYNRRS 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	1020	100.0	576	1	P02879 ricinus com
2	930.5	91.2	564	1	P06750 ricinus com
3	353.5	34.7	562	1	P28590 abrus prec
4	345	33.8	527	1	P06077 abrus prec
5	342	33.5	289	1	P09989 trichosan
6	341.5	33.5	282	1	P8184 bryonia dio
7	338	33.1	289	1	P24478 trichosan
8	330.5	32.4	563	1	P33183 sambucus ni
9	328.5	32.3	528	1	P31140 abrus prec
10	312	30.6	290	1	P33185 bryonia dio
11	300.5	29.5	250	1	P22851 luffa cylin
12	296.5	29.1	286	1	P09146 cucumis fig
13	286	28.0	254	1	P81446 viscum albu
14	284.5	27.9	277	1	P00465 luffa cylin
15	279	27.4	286	1	P29339 momordica b
16	277	27.2	286	1	P16094 momordica c
17	260	25.5	294	1	P56626 trichosan
18	242.5	23.8	316	1	P33186 gelonium mu
19	197.5	19.4	294	1	P00346 phytoacca
20	182	17.8	313	1	P10297 phytoacca
21	180.5	17.7	261	1	P23333 phytoacca
22	178	17.5	278	1	P23326 mirabilis j
23	140	13.7	280	1	P04399 hordeum vul
24	133	13.0	283	1	P22244 hordeum vul
25	131	12.8	250	1	P41391 saponaria o
26	127	12.5	253	1	P41389 saponaria o
27	124	12.2	299	1	P20656 saponaria o
28	123.5	12.1	293	1	P24476 dianthus ca
29	122	12.0	310	1	P40772 phytoacca
30	121	11.9	292	1	P27559 saponaria o
31	111	10.9	319	1	P09385 bacterioph
32	110	10.8	236	1	P27560 saponaria o
33	109.5	10.7	300	1	P25891 zea mays (m

34	107.5	10.5	301	1	RIPX_MAIZE	P28522 zea mays (m
35	103	10.1	304	1	RIP_MAIZE	P25892 zea mays (m
36	99	9.7	560	1	J160_HORVU	P00531 hordeum vul
37	90	8.8	315	1	SLTA_BP30	P08026 bacterioph
38	90	8.8	315	1	SLTA_BP30	P10149 bacterioph
39	85.5	8.4	1165	1	SVY_AQUAE	P06741 aquifex aeo
40	85	8.3	609	1	FTSH_TREPA	P083746 treponema p
41	82.5	8.1	492	1	AERA_AERTR	P09166 aeromonas t
42	81	7.9	305	1	V348_MYCPN	P75255 mycoplasma
43	78.5	7.7	485	1	AERS_AERHY	P06306 aeromonas h
44	78.5	7.7	493	1	AERA_AERHY	P09167 aeromonas h
45	76	7.5	157	1	RIP4_SAPOF	P27561 saponaria o

## ALIGNMENTS

RESULT 1	ID	RICI_RICCO	STANDARD:	PRT:	576 AA.
AC	P02879	P02880:			
DT	21-JUL-1986	(Rel. 01, Created)			
DT	13-AUG-1987	(Rel. 05, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Ricin precursor [Contains: Ricin A chain (RRNA N-glycosidase)				
DE	(EC 3.2.2.22): Ricin B chain).				
OS	Ricinus communis (Castor bean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosid 1; Malpighiales; Euphorbiaceae; Ricinus.				
OX	NCBI_TaxID=3986;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=86067214; PubMed=2999712;				
RA	Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,				
RA	Weaver R.F.;				
RT	"Genomic cloning and characterization of a ricin gene from Ricinus				
RT	communis.";				
RL	Nucleic Acids Res. 13:8019-8033(1985).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92163016; PubMed=1371405;				
RA	Tregear J.W., Roberts L.M.;				
RT	"The lectin gene family of Ricinus communis: cloning of a functional				
RT	ricin gene and three lectin pseudogenes.";				
RL	Plant Mol. Biol. 18:515-525(1992).				
RN	[3]				
RP	SEQUENCE OF 12-576 FROM N.A.				
RX	MEDLINE=85179479; PubMed=383723;				
RA	Lamb A., Roberts L.M., Lord J.M.;				
RT	"Nucleotide sequence of cloned cDNA coding for preproricin.";				
RL	Eur. J. Biochem. 148:265-270(1985).				
RN	[4]				
RP	SEQUENCE OF 36-302.				
RA	Yoshitake S., Funatsu G., Funatsu M.;				
RT	"Isolation and sequences of peptic peptides, and the complete				
RT	sequence of 11e chain of ricin-D.";				
RL	Agric. Biol. Chem. 42:1267-1274(1978).				
RN	[5]				
RP	SEQUENCE OF 315-576.				
RA	Funatsu G., Kimura M., Funatsu M.;				
RT	"Primary structure of Ala chain of ricin D.";				
RL	Agric. Biol. Chem. 43:2221-2224(1979).				
RN	[6]				
RP	CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.				
RX	MEDLINE=90344223; PubMed=1368517;				
RA	Kimura Y., Kusouku H., Tada M., Takagi S., Funatsu G.;				
RT	"Structural analyses of sugar chains from ricin A-chain variant.";				
RL	Agric. Biol. Chem. 54:157-162(1990).				
RN	[7]				
RP	REVIEW.				
RX	MEDLINE=21480122; PubMed=11595634;				
RA	Olsnes S., Kozlov J.V.;				

RT "Ricin.";  
 RL Toxicon 39:1723-1728(2001).  
 RN (8)  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=87165983; PubMed=3558397;  
 RA Monfort W., Villafraña J.E., Monzingo A.F., Ernst S.R., Katzlin B.,  
 RT Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.;  
 RL "The three-dimensional structure of ricin at 2.8 Å.";  
 J. Biol. Chem. 262:5398-5403(1987).  
 RN (9)  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=91352004; PubMed=1881881;  
 RA Katzlin B.J., Collins E.J., Robertus J.D.;  
 RT "Structure of ricin A-chain at 2.5 Å.";  
 RL Proteins 10:251-259(1991).  
 RN (10)  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.  
 RX MEDLINE=91352005; PubMed=1881882;  
 RA Rutenber E., Robertus J.D.;  
 RT "Structure of ricin B-chain at 2.5-Å resolution.";  
 RL Proteins 10:260-269(1991).  
 RN (11)  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=95082010; PubMed=7990130;  
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,  
 RT Paupit R.A.;  
 RL "X-ray structure of recombinant ricin A-chain at 1.8-Å resolution.";  
 J. Mol. Biol. 244:410-422(1994).  
 RN (12)  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.  
 RX MEDLINE=96374222; PubMed=8780513;  
 RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,  
 RT Molina-Synth M.C., Robertus J.D.;  
 RL "Structure and activity of an active site substitution of ricin A chain.";  
 Biochemistry 35:11098-11103(1996).  
 RN (13)  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=97240820; PubMed=9086280;  
 RA Van X., Hollis T., Synth M., Day P., Monzingo A.F., Milne G.W.,  
 RT Robertus J.D.;  
 RL "Structure-based identification of a ricin inhibitor.";  
 J. Mol. Biol. 266:1043-1049(1997).  
 RN (14)  
 RP MUTAGENESIS.  
 RX MEDLINE=93165632; PubMed=1287657;  
 RA Kin Y., Robertus J.D.;  
 RT "Analysis of several key active site residues of ricin A chain by mutagenesis and X-ray crystallography.";  
 Protein Eng. 5:775-779(1992).  
 CC -1- FUNCTION: Ricin is highly toxic to animal cells and to a less extent to plant cells. The A chain is responsible for inhibiting protein synthesis through the catalytic inactivation of 60S ribosomal subunits. It acts as a glycosylase that removes a specific adenine residue from an exposed loop of 28S ribosomal RNA. As this loop is involved in the binding of elongation factors, the modified ribosomes are unable to support protein synthesis. The A chain can inactivate a few thousand ribosomes per minute, thus inactivating them faster than the cell can make new ones. A single A-chain molecule can therefore kill an animal cell. The B chain binds to cell receptors and facilitates the entry into the cell of the A chain; B chains are also responsible for cell agglutination (lectin activity). It binds to beta-D-galactopyranoside moieties.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: Disulfide-linked dimer of A and B chains.  
 CC -1- DOMAIN: The B chain is composed of two domains, each domain consists of 3 homologous subdomains (alpha, beta, gamma).  
 CC -1- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.

CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -1- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).  
 CC -1- DATABASE: NME-Protein Spotlight;  
 CC WWW=Issue 31 of February 2003;  
 CC WWW="http://www.expasy.org/spotlight/articles/spl1c031.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X03179; CA265939.1; -;  
 DR EMBL; X52908; CA37095.1; -;  
 DR EMBL; X02388; CA26230.1; -;  
 DR EMBL; A12892; CA01058.1; -;  
 DR PIR; A24041; RLCSO.  
 DR PDB; 2AAT; 31-JAN-94.  
 DR PDB; 1APG; 31-JAN-94.  
 DR PDB; 1FME; 31-OCT-93.  
 DR PDB; 1IFS; 14-JAN-98.  
 DR PDB; 1IFU; 14-JAN-98.  
 DR PDB; 1RTG; 31-OCT-93.  
 DR PDB; 1OB5; 16-JUN-97.  
 DR PDB; 1OBT; 16-JUN-97.  
 DR PDB; 1BR5; 02-SEP-98.  
 DR PDB; 1BR6; 02-SEP-98.  
 DR PDB; 1IL3; 16-JAN-02.  
 DR PDB; 1IL4; 16-JAN-02.  
 DR PDB; 1IL9; 16-JAN-02.  
 DR GLYCOSULEDB; P02879; -;  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RLP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RLP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIG\_RICIN; 1.  
 DR KMW Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein; Lectin; Signal; 3D-structure.  
 FT SIGNAL 1 35  
 FT CHAIN 36 302  
 FT PEPTIDE 303 314  
 FT CHAIN 315 576  
 FT DOMAIN 321 448  
 FT DOMAIN 451 575  
 FT REPEAT 331 373  
 FT REPEAT 374 414  
 FT REPEAT 417 449  
 FT REPEAT 462 497  
 FT REPEAT 501 540  
 FT REPEAT 543 570  
 FT REPEAT 512 570  
 FT ACT\_SITE 212 212  
 FT DISULFID 294 318  
 FT DISULFID 334 353  
 FT DISULFID 377 394  
 FT DISULFID 465 478  
 FT DISULFID 504 521  
 FT CARBOHYD 45 45  
 FT CARBOHYD 271 271  
 FT CARBOHYD 409 409  
 FT CARBOHYD 449 449  
 FT CONFLICT 76 76  
 FT CONFLICT 551 551  
 FT STRAND 43 47  
 FT TURN 49 50  
 N-LINKED (GLCNAC. . .).  
 /FTid-CAR.000080.  
 N-LINKED (GLCNAC. . .) (IN MINOR FORM).  
 /FTid-CAR.000081.  
 N-LINKED (GLCNAC. . .).  
 N-LINKED (GLCNAC. . .).  
 E -> D (IN REF. 3).  
 A -> R (IN REF. 3).



Query Match 100.0%; Score 1020; DB 1; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-87;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFPKYPIINFTTACATVOSTYTNFIRAVRGRLTTGADVRIHEIPVLPNRVGLPIINORFIV 60  
 |||||||  
 DB 36 IFFPKYPIINFTTACATVOSTYTNFIRAVRGRLTTGADVRIHEIPVLPNRVGLPIINORFIV 95

QY 61 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 120  
 |||||||  
 DB 96 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 155

QY 121 GNYDRLEQLAGNLENIELGNGPLEAISALYYSTGCTQLPTLARSFFICQIMISEAR 180  
 |||||||  
 DB 156 GNYDRLEQLAGNLENIELGNGPLEAISALYYSTGCTQLPTLARSFFICQIMISEAR 215

QY 181 FOYTEGEMRTIRIRYNNRS 198  
 |||||||  
 DB 216 FOYTEGEMRTIRIRYNNRS 233

RESULT 2  
 AGGL\_RICCO STANDARD; PRT; 564 AA.  
 ID AGGL\_RICCO STANDARD; PRT; 564 AA.  
 AC P06730;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Agglutinin B chain].  
 OS Ricinus communis (Castor bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
 NCBI\_TaxID=3986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86059449; PubMed=2999130;  
 RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;  
 RT "The primary sequence of Ricinus communis agglutinin. Comparison with ricin.";  
 RL J. Biol. Chem. 260:15682-15686(1985).  
 RN [2]  
 RP SEQUENCE OF 303-564.  
 RC TISSUE=Seed;  
 RA Ataki T., Yoshioaka Y., Funatsu G.;  
 RT "The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds.";  
 RL Biochim. Biophys. Acta 872:277-285(1986).  
 RN [3]  
 RP SEQUENCE OF 303-337.  
 RX MEDLINE=80178723; PubMed=6768555;  
 RA Lin T.T.-S., Li S.S.-L.;  
 RT "Purification and physicochemical properties of ricins and agglutinins from Ricinus communis.";  
 RL Eur. J. Biochem. 105:453-459(1980).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M12089; AAA33869.1; -  
 DR EMBL: S40368; AAB22584.1; -

DR PIR: A24261; RLC5AG.  
 DR HSSP: P02879; IIR6.  
 DR GlycosultedB; P06750; -.  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHICARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS0231; RICIN\_B-LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Hydrolyase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein; Lectin; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25  
 FT PROPEP 251 302  
 FT CHAIN 303 564  
 FT DOMAIN 309 436  
 FT REPEAT 439 563  
 FT REPEAT 319 361  
 FT REPEAT 362 402  
 FT REPEAT 405 437  
 FT REPEAT 450 485  
 FT REPEAT 489 528  
 FT REPEAT 531 558  
 FT ACT\_SITE 200 200  
 FT DISULFID 282 306  
 FT DISULFID 322 341  
 FT DISULFID 365 382  
 FT DISULFID 453 466  
 FT DISULFID 492 509  
 FT CARBOHYD 34 34  
 FT CARBOHYD 259 259  
 FT CARBOHYD 397 397  
 FT CARBOHYD 437 437  
 FT CONFLICT 331 331  
 FT CONFLICT 362 362  
 FT CONFLICT 374 374  
 FT CONFLICT 404 404  
 FT CONFLICT 552 552  
 SQ SEQUENCE 564 AA: 62851 MW: 14552727609759 CRC64;

Query Match 91.2%; Score 930.5; DB 1; Length 564;  
 Best Local Similarity 91.9%; Pred. No. 5e-79;  
 Matches 182; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 1 IFFPKYPIINFTTACATVOSTYTNFIRAVRGRLTTGADVRIHEIPVLPNRVGLPIINORFIV 60  
 |||||||  
 DB 25 IFFPKYPIINFTTACATVOSTYTNFIRAVRGRLTTGADVRIHEIPVLPNRVGLPIINORFIV 84

QY 61 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 120  
 |||||||  
 DB 85 ELSNHAELSVTLADVTNAVYVGRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 144

QY 121 GNYDRLEQLAGNLENIELGNGPLEAISALYYSTGCTQLPTLARSFFICQIMISEAR 180  
 |||||||  
 DB 145 GNYDRLEQLAGNLENIELGNGPLEAISALYYSTGCTQLPTLARSFFICQIMISEAR 203

QY 181 FOYTEGEMRTIRIRYNNRS 198  
 |||||||  
 DB 204 FOYTEGEMRTIRIRYNNRS 221

RESULT 3  
 ABRC\_ABRPR STANDARD; PRT; 562 AA.  
 ID ABRC\_ABRPR STANDARD; PRT; 562 AA.  
 AC P28590;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Abirin-c precursor [Contains: Abirin-c A chain (rRNA N-glycosidase) (EC 3.2.2.22); Abirin-c B chain].  
 OS Abrus precatorius (Indian liquorice) (Crab's eye).

CC	Enkaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
CC	- Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
CC	eurosid_1; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
CC	NCBI_TaxID=3816;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Leaf;
RC	MEDLINE=91266957; PubMed=2050149;
RA	Wood K.A., Lord J.M., Mawrzynszak E.J., Platak M.;
RT	"Preproabrin: genomic cloning, characterisation and the expression of
RT	the A-chain in Escherichia coli.";
RL	Eur. J. Biochem. 196;723-732(1991).
CC	- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC	SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC	SUBUNTS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE
CC	B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE
CC	BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.
CC	- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC	specific adenosine on the 28S rRNA.
CC	- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC	- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC	CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC	- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
CC	INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
CC	- SIMILARITY: Contains 2 ricin B-type lectin domains.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; X55667; CAA39202.1; .
DR	PIR; S16022; S16022.
DR	HSSP: P11140; IABR.
DR	InterPro; IPR000772; Ricin_B_lectin.
DR	InterPro; IPR001574; RIP.
DR	Pfam; PF00652; Ricin_B_lectin; 6.
DR	Pfam; PF00161; RIP; 1.
DR	PRINTS; PR00396; SHIGARICIN.
DR	SMART; SM00458; RICIN; 2.
DR	PROSITE; PS00231; RICIN_B_LECTIN; 2.
DR	PROSITE; PS00275; SHIGA_RICIN; 1.
KW	Plant defense; Hydrolyase; Protein synthesis inhibitor; Toxin; Repeat;
KW	Glycoprotein; Lectin; Signal; Pyrrolidone carboxylic acid.
FT	SIGNAL 1 34
FT	CHAIN 35 285
FT	PEPTIDE 286 295
FT	CHAIN 286 562
FT	DOMAIN 307 434
FT	DOMAIN 437 561
FT	REPEAT 317 359
FT	REPEAT 360 400
FT	REPEAT 403 435
FT	REPEAT 448 483
FT	REPEAT 487 526
FT	REPEAT 529 562
FT	ACT_SITE 198 198
FT	DISULFID 281 303
FT	DISULFID 320 339
FT	DISULFID 363 380
FT	DISULFID 451 464
FT	DISULFID 490 507
FT	MOD_RES 35 35
FT	CARBOHYD 234 234
FT	CARBOHYD 395 395
FT	CARBOHYD 435 435
SEQ	SEQUENCE 562 AA; 62817 MW; 1PDDABC7D/BA6278 CR664;

Query Match 34.7%; Score 353.5; DB 1; Length 562;

	Best Local Similarity	45.5%; Pred.	No. 2.6e-25;	
	Matches	86;	Conservative	24; Mismatches 70; Indels 9; Gaps 4.
QY	5 QYPIINETAGTQSYSTNFTIRAVRGRLTGADVRHPIPLPNRGVGPINQSEFIVEELSN	64		
Db	35 QDQVKEFTTEGKISQSYKQFIEALRQRLTGS--LIHDIPLVPPPTVEERNKRYIYELSN	92		
QY	65 HAEIYVTLALDYTNAYVVGVRAGNSANVFEPHDNDEDAEALTHLFTDYQNRYPFARGAND	124		
Db	93 SERKESELEVDIDYTNAYVVAVRAGSOSYFL--KDAFASASTYLPFGTO-RYSLRFGSGTG	148		
QY	125 RLEQLGNIGRENIIEGNGPLLEBAISALTYSTGTGLOPLIARSPFIICMIQISEAARFOYI	184		
Db	149 DIERMAHQRRREISIGLQALTHAIS--FLRSGASNDDEKARFLIYIOWASEANAYRYI	205		
QY	185 EGEMLRTIR	193		
Db	206 SNRQGVSIK	214		

```

CC AC ID ABRB_ABRPR STANDARD; PRT; 527 AA.
CC DT Q06077; P81374;
CC DT 15-DEC-1998 (Rel. 37, Created)
CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DT Abrin-b precursor [contains Abrin-b A chain (RNA N-glycosidase)
CC DE (EC 3.2.2.22); Abrin-b B chain].
CC DE Abrus precatorius (Indian licorice) (Crab's eye).
CC OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC OC eucotsids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
CC OX NCBI_TaxId=3816;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=93132798; PubMed=8421313;
CC RA Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;
CC RT "Primary structure of three distinct isoabrin determined by cDNA
CC RT sequencing. Conservation and significance.";
CC RL J. Mol. Biol. 229:263-267(1993).
CC RN [2]
CC RP SEQUENCE OF 260-527.
CC RC TISSUE=Seed;
CC RX MEDLINE=93169023; PubMed=7763422;
CC RA Kimura M., Sumizawa T., Funatsu G.;
CC RT "The complete amino acid sequences of the B-chains of abrin-a and
CC RL abrin-b, toxic proteins from the seeds of Abrus precatorius.";
CC RL Biosci. Biotechnol. Biochem. 57:166-169(1993)
CC CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC CC SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC CC ABRIN-A IS MORE TOXIC THAN RICIN.
CC CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
CC CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
CC CC PRECEDES ENDOCYTOSIS.
CC CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC CC specific adenosine on the 28S RNA.
CC CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
CC CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC CC or send an email to license@isb-slb.ch).
CC CC -----

```

```

DR EMBL: M96345; AAA32625.1; -.
DR PIR: S32430; S32430.
DR HSSP: P11140; 1ABR.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS50231; RICIN_B_LLECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein; Lectin; Pyroglutamate carboxylic acid.
KM CHAIN 1
FT CHAIN 1 250
FT PEPTIDE 251 260
FT CHAIN 251 527
FT CHAIN 251 527
FT DOMAIN 272 399
FT DOMAIN 402 526
FT DOMAIN 282 324
FT REPEAT 325 365
FT REPEAT 368 400
FT REPEAT 413 448
FT REPEAT 452 491
FT REPEAT 494 527
FT ACT_SITE 163 165
FT DISULFID 246 268
FT DISULFID 285 304
FT DISULFID 328 345
FT DISULFID 416 429
FT DISULFID 455 472
FT MOD_RES 1 1
FT CARBOHYD 110 110
FT CARBOHYD 360 360
FT CARBOHYD 400 400
FT CONFLICT 282 282
FT CONFLICT 291 291
FT CONFLICT 350 351
FT CONFLICT 378 378
FT CONFLICT 426 426
FT CONFLICT 428 428
FT CONFLICT 431 431
FT CONFLICT 484 484
FT CONFLICT 491 491
FT CONFLICT 493 493
FT CONFLICT 502 502
FT CONFLICT 509 509
FT CONFLICT 513 513
FT CONFLICT 516 516
SO SEQUENCE 527 AA; 59114 MW; 3253AE90CE9494A CRC64;

Query Match 33.8%; Score 345; DB 1; Length 527;
Best Local Similarity 45.5%; Pred. No. 1.5e-24;
Matches 87; Conservative 22; Mismatches 72; Indels 10; Gaps 4;

OY 5 QPPIINFTAGATVOSTYTFIAVAGRLTTGADVADHEIPVLPNVRGLPINOEFILVELSN 64
DB 1 QDVQIKFTTEGATSSQYKQIFALRQRLTGG--LHGGIPVLPDPTTLQERNRYISVELSN 58
OY 65 HAEIYSTLADVTNVAVGVYRAGNSAVFPHNOADAETPLFTDYNQRTFAAGVND 124
DB 59 STTESLEAGIDVSNVAIVRAGNSYFL--RDAPTSASRLFTGTGTO-QYSLRNGSYIT 114
OY 125 RLEQIAGNRENIENLENGPLEEASLYYSTGTGTPLTARSLFFICIQIMSEARFOYI 184
DB 115 DIERLAROTRQOIPLEGIALRHAISFL---QSGTDDEIATRLTVIIOQMSAARYMFI 170
OY 185 EGEWRTIRYN 195
DB 171 SYRGVSIKTN 181

```

```

RIP1 TRIKI
ID RIP1 TRIKI STANDARD: PRT: 289 AA.
AC P09689;
DT 01-MAR-1989 (rel. 10, Created)
DT 01-NOV-1990 (rel. 16, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE Ribosome-inactivating protein alpha-trichosanthin precursor
DE (RNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS)
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OC NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Maximowicz; TISSUE=leaf;
RX MEDLINE=91153657; PubMed=199291;
RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.R., Yeung H.W.;
RT "Cloning of trichosanthin cDNA and its expression in Escherichia
coll.";
RL Gene 97:267-272(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Maximowicz; TISSUE=leaf;
RX MEDLINE=90256790; PubMed=2341400;
RA Chow T., Feldman R.A., Lovett M., Platak M.;
RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a
type I ribosome-inactivating protein.";
RL J. Biol. Chem. 265:8670-8674(1990).
RN [3]
RP SEQUENCE OF 24-270.
RC STRAIN-Maximowicz; TISSUE=tuberous root;
RX MEDLINE=90256789; PubMed=2341399;
RA Collins E.J., Roberius J.D., Lopresti M., Stone K.L., Williams K.R.,
Wu P., Hwang K., Platak M.;
RT "Primary amino acid sequence of alpha-trichosanthin and molecular
models for adrin A-chain and alpha-trichosanthin.";
RL J. Biol. Chem. 265:8665-8669(1990).
RN [4]
RP SEQUENCE OF 24-270.
RC TISSUE=tuberous root;
RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,
Tian G.Y., Ni C.Z.;
RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and
application.";
RL Pure Appl. Chem. 58:789-798(1986).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).
RX MEDLINE=94344957; PubMed=8066085;
RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;
RT "Structure of trichosanthin at 1.88-A resolution.";
RL Proteins 19:4-13(1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX MEDLINE=95344383; PubMed=7619070;
RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active-centre geometry and
deputinating mechanism of two ribosome-inactivating proteins.";
RL Biochem. J. 309:285-298(1995).
RN [7]
RP FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS
CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT
INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.
RN [8]
RP CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
specific adenosine on the 28S rRNA.
RN [9]
RP SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
RN [10]
RP TYPE 1 RIP SUBFAMILY.

```

RESULT 5



RT MEDLINE=94003077; PubMed=8400135;  
RA Gribes T., Cflores L., Ferreras J.M., Rojo M.A., Iglesias R.,  
R Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;  
RT Isolation and partial characterization of nigrin b, a non-toxic  
RT novel type 2 ribosome-inactivating protein from the bark of *Sambucus*  
RT nigra L.<sup>1</sup>;  
RL Plant Mol. Biol. 22:1181-1186(1993).  
CC -1- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN







```

Db      63 STEIGDITNAYVAAYAGTGYFLRAPSASD-----YLFTGT-DQHLPEYGYGDL 116
Qy      127 EQLAGNLENIEGNGPLFEALYYSTGGTQLPTLARSFFICQIMISEARFOYIEG 186
Db      117 ERMAHQSRQQLPLGLALTLHGIS---PFRSGNDNEKARTLLYIOMVAEARFRYSN 173
Qy      187 EMRTTR 193
Db      174 RVRVSIQ 180

RESULT 10
R1PL_BRYDI STANDARD; PRT: 290 AA.
AC P33185; Q9S819;
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein bryodin I precursor (rRNA N-glycosidase)
DE (EC 3.2.2.22) (BDL)
OS Bryonia dioica (Red bryony).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.
OX NCBI_TaxID=3652;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RC TISSUE=Leaf;
RX MEDLINE=97228081; PubMed=9115985;
RA Siegall C.B.; Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,
RA Siegall C.B.;
RT "Molecular, biological, and preliminary structural analysis of
RT recombinant bryodin I, a ribosome-inactivating protein from the plant
RT Bryonia dioica";
RL Biochemistry 36:3095-3103(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Siegall C.B.;
RT "Cloning and expression of a gene encoding bryodin I from Bryonia
RT dioica.";
RL Patent number US5541110, 30-JUL-1996.
RN [3]
RP SEQUENCE OF 24-66.
RC TISSUE=Seed;
RX MEDLINE=89326691; PubMed=2753596;
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
RA Lapi D.;
RT "N-terminal sequence of some ribosome-inactivating proteins.";
RL Int. J. Pept. Protein Res. 33:263-267(1989).
RN [4]
RP SEQUENCE OF 24-43.
RC TISSUE=Root;
RX MEDLINE=95151812; PubMed=7849072;
RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixon B.,
RA Marquardt H.;
RT "Characterization of ribosome-inactivating proteins isolated from
RT Bryonia dioica and their utility as carcinoma-reactive
RT immunoonjugates.";
RL Bioconj. Chem. 5:423-429(1994).
CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS
CC PROTEIN SYNTHESIS IN ANIMAL CELLS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO
CC PRODUCE A SHORTER PROTEIN.
CC -1- BIOTECHNOLOGY: Especially useful as immunotoxin for
CC pharmacological applications as it has low toxicity in rats and
CC mice but is potent once inside target cells.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J24020; NOT_ANNOTATED_CDS.
DR PIR: S16491; S16491.
DR PDB: 1BR7; 04-MAR-98.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP. 1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
KW 3D-structure; Multigene family; Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 270
FT PROPEP 271 290
FT ACT_SITE 183 183
FT ACT_SITE 212 212
FT CARBOHYD 214 214
FT CARBOHYD 250 250
FT MOTAGEN 212 212
FT CONFLICT 61 65
FT STRAND 25 28
FT TURN 30 31
FT TURN 34 31
FT HELIX 47 47
FT STRAND 50 54
FT TURN 55 55
FT STRAND 57 60
FT HELIX 66 69
FT STRAND 70 76
FT TURN 78 79
FT STRAND 82 88
FT TURN 89 92
FT STRAND 93 93
FT TURN 100 101
FT STRAND 102 105
FT HELIX 109 114
FT TURN 115 117
FT TURN 120 121
FT STRAND 124 127
FT HELIX 134 141
FT TURN 142 142
FT HELIX 145 147
FT STRAND 150 150
FT HELIX 152 163
FT TURN 164 165
FT HELIX 167 186
FT STRAND 187 187
FT HELIX 188 196
FT STRAND 202 202
FT HELIX 206 213
FT TURN 214 214
FT HELIX 215 225
FT TURN 226 230
FT STRAND 231 239
FT TURN 241 242
FT STRAND 245 250
FT TURN 251 252
FT HELIX 254 257
FT TURN 258 259
FT STRAND 260 260
FT TURN 263 263
FT HELIX 266 268
SQ SEQUENCE 290 AA; 31788 MW; E966CD9C031A42DB CRC64;

Query Match 30.6%; Score 312; DB 1; Length 290;
Best Local Similarity 37.3%; Pred. No. 8, 2e-22;
Matches 69; Conservative 46; Mismatches 58; Indels 12; Gaps 5;

```



[illegible]

MLA\_VISAL  
ID MLA\_VISAL STANDARD; PRT; 254 AA.  
AC P01446;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Beta-galactoside specific lectin I A chain (MLA) (ML-I A) (rRNA N-glycosidase) (EC 3.2.2.22).  
OS Viscum album (European mistletoe).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Santalales; Santalaceae; Viscum.  
OX NCBI\_TaxID=3972;  
RN [1]  
RP SEQUENCE  
RC STRAIN=Subsp. album;  
RX MEDLINE=97134581; PubMed=8980141;  
RA Soler M.H., Stoeva S., Schwamborn C., Wilhelm S., Stiefel T., Voelter W.;  
RT "Complete amino acid sequence of the A chain of mistletoe lectin I."; FBBS Lett. 399:153-157(1996).  
RL  
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE B CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY INTO THE CELL OF THE A CHAIN; B CHAINS ARE ALSO RESPONSIBLE FOR CELL AGGLUTINATION (LECTIN ACTIVITY).  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
CC -1- SUBUNIT: Disulfide-linked dimer of A and B chains.  
CC -1- PHARMACUTICAL: Due to its immunomodulatory effects it is being studied in clinical trials in cancer patients as it may slow the growth of cancer cells and be an effective treatment for solid tumors.  
CC -1- MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA AND NON-GLYCOSYLATED FORM MLA.  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
DR PIR: PD0018; PD0018.  
DR HSSP: P11140; 1ABR.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR PROSITE: PS00275; SHIGARICIN; FALSE\_NEG.  
KW Plant defense; Hydrolyase; Toxin; Repeat; Glycoprotein; Lectin.  
FT ACT\_SITE 165 165  
FT CARBOXYD 112 112  
FT VARIANT 15 15  
FT VARIANT 66 66  
FT VARIANT 112 112  
FT VARIANT 116 116  
FT VARIANT 133 134  
FT VARIANT 140 140  
FT VARIANT 144 144  
FT VARIANT 151 151  
FT VARIANT 179 179  
FT VARIANT 184 184  
FT VARIANT 190 190  
FT VARIANT 218 218  
FT VARIANT 223 224  
FT VARIANT 231 231  
FT VARIANT 235 235  
SQ SEQUENCE 254 AA; 28478 MW; 53BAF98D3E0FE67 CRC64;  
Query Match 28.0%; Score 286; DB 1; Length 254;  
Best Local Similarity 38.4%; Pred. No. 1.8e-19;  
Matches 81; Conservative 36; Mismatches 36; Indels 38; Gaps 10;

DB 67 TALAIDVTNAIVATGAGOSTFLR-DAPKAE--THLEGT-TRSSLPGNSIPDLERYA 122  
OY 131 GNLRRENIEGNGPLEEASALYYSTGTQDPTLARSFICIONISEARF-----Q 182  
DB 123 GH-RQPIPLGIDQLIGSTALRF--PGSRTGQASILLIQMISEARFNPLMRYRQ 178  
OY 183 YIE-----GEMRTIRYN 195  
DB 179 YINSGASLPDVYMLEETSMGQOSTOVQHS 209  
RESULT 14  
RIPA\_LURCY  
ID RIPA\_LURCY STANDARD; PRT; 277 AA.  
AC 000465;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ribosome-inactivating protein luffin-alpha precursor (rRNA N-glycosidase) (EC 3.2.2.22).  
OS Luffa cylindrica (Smooth loofah) (Sponge gourd).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OX eucosids I; Cucurbitales; Cucurbitaceae; Luffa.  
OX NCBI\_TaxID=3670;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Seed;  
RX MEDLINE=92288316; PubMed=1600156;  
RA Katsoka J., Habuka N., Miyano M., Masuta C., Koizumi A.;  
RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-inactivating protein from Luffa cylindrica.";  
RL Plant Mol. Biol. 18:1199-1202(1992).  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 1 RIP SUBFAMILY.  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to [license@isb.ch](mailto:license@isb.ch)).  
CC  
DR EMBL: X62371; CAA44229.1; -.  
DR PIR: S22494; S22494.  
DR HSSP: P16094; 1AHC.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR PROSITE: PS00275; SHIGARICIN; 1.  
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolyase;  
FT TOXIN; SIGNAL 1 19  
FT CHAIN 20 277  
FT SIGNAL 1 19  
FT ACT\_SITE 179 179  
FT SEQUENCE 277 AA; 30212 MW; EA17FC27998C25AC CRC64;  
Query Match 27.9%; Score 284.5; DB 1; Length 277;  
Best Local Similarity 33.5%; Pred. No. 2.8e-19;  
Matches 63; Conservative 45; Mismatches 69; Indels 11; Gaps 3;

QY 129 LAGNLENIENLGNPLLEAISAALYYSTGCTQLPTLARSFLICIQMISEARFOYIEGEM 188  
 Db 136 AAGKIREKIPGLFALPSAITTLTFYNA-----QAPSAALVLIQTAAEARFKYIERHV 192  
 QY 189 RTIRIRYN 196  
 Db 191 IERISKNO 198

## RESULT 15

RIP2\_MOMBA

ID RIP2\_MOMBA STANDARD; PRT; 286 AA.

AC P29339;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ribosome-inactivating protein momordin II precursor (rRNA  
 N-glycosidase) (EC 3.2.2.22).  
 OS Momordica balsamina (Bitter melon) (Balsam pear).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids I; Cucurbitales; Cucurbitaceae; Momordica.  
 OC NCBI\_TaxID=3672;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seed;  
 RX MEDLINE=93027170; PubMed=1408771;  
 RA Ortigao M., Better M.;  
 RT "Momordin II, a ribosome inactivating protein from Momordica  
 balsamina, is homologous to other plant proteins.";  
 RL Nucleic Acids Res. 20:4662-4662(1992).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: Z12175; CAA78166.1; -  
 DR PIR: S25560; S25560.  
 DR PDB: 1CF5; 07-JUN-99.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; signal;  
 KW 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 286 RIBOSOME-INACTIVATING PROTEIN MOMORDIN  
 FT II.  
 FT ACT\_SITE 181 181 BY SIMILARITY.  
 SQ SEQUENCE 286 AA; 32031 MW; 3B89FF1AE6B25986 CRC64;

Query Match 27.4%; Score 279; DB 1; Length 286;  
 Best Local Similarity 35.3%; Pred. No. 9.5e-19;  
 Matches 66; Conservative 42; Mismatches 67; Indels 12; Gaps 5;

QY 9 INETAGATVOSTYNEIRAVRGLTGTGADVREHIEIPVLRVGLPINORFIIELSNHAE 68  
 Db 25 VNEDLSTATKTYTKFIEDERATLPESHKV-YDIPLYSTIS--DSRRFILDITSYAE 81  
 QY 69 SYTLADYTNNAVYVGRAGNSAYFFHPDQEDAEATLFTDVONRYTFAGGNDRLQ 128  
 Db 82 TTSVADYTNNAVYVGRAGNSAYFF--KESPEAYNILEFKGTR-KITLPYTGNYENLOT 137  
 QY 129 LAGNLENIENLGNPLLEAISAALYYSTGCTQLPTLARSFLICIQMISEARFOYIEGEM 188

Db 138 AAGKIREKIPGLFALPSAITTLTFYNA-----QAPSAALVLIQTAAEARFKYIERHV 192  
 QY 189 RTIRIRYN 195  
 Db 193 AKYVATN 199

Search completed: September 16, 2003, 11:46:03  
 Job time : 7.94128 secs

1  
2  
3  
4

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:40:08 ; Search time 31.5828 Seconds  
(without alignments)  
1617.791 Million cell updates/sec

Title: US-10-083-336a-3

Perfect score: 1020

Sequence: 1 IFFKQYPIINFTTAGATVQS.....ARFOYIEGEMPTRIYNNRS 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP:archaea:\*  
2: SP:bacteria:\*  
3: SP:fungi:\*  
4: SP:human:\*  
5: SP:invertebrate:\*  
6: SP:mammal:\*  
7: SP:mhc:\*  
8: SP:organelle:\*  
9: SP:phage:\*  
10: SP:plant:\*  
11: SP:rodent:\*  
12: SP:virus:\*  
13: SP:vertebrate:\*  
14: SP:unclassified:\*  
15: SP:ivirius:\*  
16: SP:bacteriap:\*  
17: SP:archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1017	99.7	541	10	Q41174	094174 ricinus com
2	401.5	39.4	580	10	Q94BM3	094bm3 cinnaomum
3	397.5	39.0	580	10	Q94BM4	094bm4 cinnaomum
4	397.5	39.0	581	10	Q94BM5	094bm5 cinnaomum
5	395.5	38.8	549	10	Q9FV22	09fv22 cinnaomum
6	350.5	34.4	563	10	004367	004367 sambucus nl
7	347.5	34.1	528	10	Q9AVR2	09avr2 sambucus eb
8	344.5	33.8	568	10	006076	006076 abrus prece
9	340	33.3	289	10	Q94KE4	094ke4 trichosan
10	338	33.1	247	10	Q9LRE3	09lre3 trichosan
11	338	33.1	289	10	Q41216	041216 trichosan
12	336.5	33.0	252	10	Q98760	098760 abrus prece
13	330.5	32.4	563	10	Q94552	094552 sambucus nl
14	330.5	32.4	563	10	Q98732	098732 sambucus nl
15	329.5	32.3	252	10	Q98761	098761 abrus prece
16	324	31.8	270	10	Q8LPV7	08lpv7 trichosan

## ALIGNMENTS

17	323.5	31.7	251	10	Q96236	096236 abrus prece
18	322.5	31.6	251	10	Q96237	096237 abrus prece
19	319	31.3	565	10	004071	004071 sambucus nl
20	317	31.1	270	10	Q41611	041611 trichosan
21	316.5	31.0	251	10	Q96235	096235 abrus prece
22	314.5	30.8	547	10	Q9M6E9	09m6e9 abrus prece
23	310.5	30.4	566	10	004072	004072 sambucus nl
24	307.5	30.1	278	10	000980	000980 luffa cylin
25	291	28.5	570	10	Q41358	041358 sambucus nl
26	285	27.9	249	10	Q8LK05	08lk05 sambucus nl
27	285	27.9	570	10	Q22415	022415 sambucus nl
28	283.5	27.8	604	10	Q9M654	09m654 polygonat
29	283	27.7	251	10	Q8LK04	08lk04 viscum albu
30	282	27.6	254	10	Q8LK06	08lk06 viscum albu
31	280	27.5	264	10	Q9F5H2	09f5h2 momordica c
32	278	27.3	293	10	Q8S452	08s452 jatropa cu
33	277.5	27.2	565	10	Q8W243	08w243 viscum albu
34	277	27.2	286	10	Q9FUV7	09fuv7 momordica c
35	277	27.2	569	10	P93543	093543 sambucus nl
36	275	27.0	531	10	Q8RXH6	08rxh6 viscum albu
37	273	26.8	249	10	Q8RXH7	08rxh7 viscum albu
38	272	26.7	286	10	Q41257	041257 momordica c
39	272	26.7	298	10	Q04358	004358 iris hollan
40	270	26.5	592	10	Q8W2E7	08w2e7 iris hollan
41	268	26.3	300	10	004356	004356 iris hollan
42	266.5	26.1	603	10	Q9M653	09m653 polygonat
43	262.5	25.7	299	10	Q8GZN9	08gzn9 euphorbia s
44	261	25.6	293	10	Q8VYU0	08vyu0 jatropa cu
45	259	25.4	573	10	Q8W2E8	08w2e8 iris hollan

## RESULT 1

Q41174 PRELIMINARY: PRT: 541 AA.  
ID Q41174;  
AC Q41174;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Protein A chain (EC 3.2.2.22) (rRNA N-glycosidase)  
DE (Fragment).  
OS Ricinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OK NCBI\_TaxID=3968;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92338377; PubMed=1633311;  
RA Roberts L.M., Tregear J.W., Lord J.M.;  
RT "Molecular cloning of ricin".  
RL Targeted Diagn. Ther. 7:81-97(1992).  
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RNA.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL: S40366; AAR22582.1; -.  
DR HSSP: P02879; IRR6.  
DR InterPro: IPR000772; Ricin\_B\_lectin.  
DR InterPro: IPR001574; RIP.  
DR InterPro: IPR001400; Somatotropin.  
DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR SMART: SM00458; RICIN; 2.  
DR PROSITE: PS00231; RICIN\_B\_LLECTIN; 2.  
DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
DR PROSITE: PS00338; SOMATOTROPIN\_2; 1.  
KW Hydrolase; Toxin.  
FT NON\_TER 1  
SQ SEQUENCE 541 AA: 60281 MW: 28783CDEFIF2E9D9 CRC64:

Query Match 99.7%; Score 1017; DB 10; Length 541;  
 Best Local Similarity 99.5%; Pred. No. 1.9e-87;  
 Matches 197; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFKQYPIINFTAGATVOSTYNTFIRAVGRRLTGADVRHEIPVLPNRYGLPIINORFILLV 60  
 1 IFFKQYPIINFTAGATVOSTYNTFIRAVGRRLTGADVRHEIPVLPNRYGLPIINORFILLV 60  
 DB 1 IFFKQYPIINFTAGATVOSTYNTFIRAVGRRLTGADVRHEIPVLPNRYGLPIINORFILLV 60  
 QY 61 ELSNHEALSTLADVTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 120  
 61 ELSNHEALSTLADVTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 120  
 DB 61 ELSNHEALSTLADVTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 120  
 QY 121 GNDRELOLAGNRENTEIGNGLPEEAISALYYSTGCTQLPTLARSFFICIMISEAR 180  
 121 GNDRELOLAGNRENTEIGNGLPEEAISALYYSTGCTQLPTLARSFFICIMISEAR 180  
 DB 121 GNDRELOLAGNRENTEIGNGLPEEAISALYYSTGCTQLPTLARSFFICIMISEAR 180  
 QY 181 FOYEGERMTRIRYRNR 198  
 181 FOYEGERMTRIRYRNR 198  
 DB 181 FOYEGERMTRIRYRNR 198

RESULT 2  
 Q94BW3 PRELIMINARY; PRT; 580 AA.  
 AC 094BW3:  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomim III precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomim proteins and study of their expression  
 RT patterns";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC EMBL: AY039803; AAK82460.1;  
 DR InterPro: IPR000772; R1cin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; R1cin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 DR HydroLase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 580  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIM III.  
 SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7FB58 CRC64;

Query Match 39.4%; Score 401.5; DB 10; Length 580;  
 Best Local Similarity 50.0%; Pred. No. 2.8e-29;  
 Matches 94; Conservative 28; Mismatches 59; Indels 7; Gaps 5;

QY 6 YPIINFTAGATVOSTYNTFIRAVGRRLTGADVRHEIPVLPNRYGLPIINORFILLV 64  
 6 YPIINFTAGATVOSTYNTFIRAVGRRLTGADVRHEIPVLPNRYGLPIINORFILLV 64  
 DB 6 YPIINFTAGATVOSTYNTFIRAVGRRLTGADVRHEIPVLPNRYGLPIINORFILLV 64  
 QY 33 YQVTFITTKNATKSTYQFIALRAQLASGEE-PHGIFVMRERSTVPDSKRFFILVELSN 91  
 33 YQVTFITTKNATKSTYQFIALRAQLASGEE-PHGIFVMRERSTVPDSKRFFILVELSN 91  
 DB 33 YQVTFITTKNATKSTYQFIALRAQLASGEE-PHGIFVMRERSTVPDSKRFFILVELSN 91  
 QY 65 HAEISVTALADVTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 124  
 65 HAEISVTALADVTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 124  
 DB 65 HAEISVTALADVTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 124  
 QY 92 AADSPVALAVDTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 148  
 92 AADSPVALAVDTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 148  
 DB 92 AADSPVALAVDTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 148  
 QY 125 RLEQLAGNRENTEIGNGLPEEAISALYYSTGCTQLPTLARSFFICIMISEAR 184  
 125 RLEQLAGNRENTEIGNGLPEEAISALYYSTGCTQLPTLARSFFICIMISEAR 184  
 DB 125 RLEQLAGNRENTEIGNGLPEEAISALYYSTGCTQLPTLARSFFICIMISEAR 184  
 QY 149 DIERVAGELREELILGMDPLENAISALWISNL--NQORALARSLIVYQMWAEAVRFRFI 206  
 149 DIERVAGELREELILGMDPLENAISALWISNL--NQORALARSLIVYQMWAEAVRFRFI 206  
 DB 149 DIERVAGELREELILGMDPLENAISALWISNL--NQORALARSLIVYQMWAEAVRFRFI 206

QY 185 EGMRTRI 192  
 185 EGMRTRI 192  
 DB 207 EYVRESI 214

RESULT 3  
 Q94BW4 PRELIMINARY; PRT; 580 AA.  
 AC 094BW4:  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomim II precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomim proteins and study of their expression  
 RT patterns";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC EMBL: AY039802; AAK82459.1;  
 DR InterPro: IPR000772; R1cin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; R1cin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 DR HydroLase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 580  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMIM II.  
 SQ SEQUENCE 580 AA; 64265 MW; 37E4289ECC0CBFF CRC64;

Query Match 39.0%; Score 397.5; DB 10; Length 580;  
 Best Local Similarity 49.5%; Pred. No. 6.7e-29;  
 Matches 93; Conservative 30; Mismatches 58; Indels 7; Gaps 5;

QY 6 YPIINFTAGATVOSTYNTFIRAVGRRLTGADVRHEIPVLPNRYGLPIINORFILLV 64  
 6 YPIINFTAGATVOSTYNTFIRAVGRRLTGADVRHEIPVLPNRYGLPIINORFILLV 64  
 DB 6 YPIINFTAGATVOSTYNTFIRAVGRRLTGADVRHEIPVLPNRYGLPIINORFILLV 64  
 QY 33 YQVTFITTKNATKSTYQFIALRAQLASGEE-PHGIFVMRERSTVPDSKRFFILVELSN 91  
 33 YQVTFITTKNATKSTYQFIALRAQLASGEE-PHGIFVMRERSTVPDSKRFFILVELSN 91  
 DB 33 YQVTFITTKNATKSTYQFIALRAQLASGEE-PHGIFVMRERSTVPDSKRFFILVELSN 91  
 QY 65 HAEISVTALADVTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 124  
 65 HAEISVTALADVTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 124  
 DB 65 HAEISVTALADVTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 124  
 QY 92 AADSPVALAVDTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 148  
 92 AADSPVALAVDTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 148  
 DB 92 AADSPVALAVDTNATVYVGRAGNSAYFFPHPNQDEAEATHLFTDVONRTTFAG 148  
 QY 125 RLEQLAGNRENTEIGNGLPEEAISALYYSTGCTQLPTLARSFFICIMISEAR 184  
 125 RLEQLAGNRENTEIGNGLPEEAISALYYSTGCTQLPTLARSFFICIMISEAR 184  
 DB 125 RLEQLAGNRENTEIGNGLPEEAISALYYSTGCTQLPTLARSFFICIMISEAR 184  
 QY 149 DIERVAGELREELILGMDPLENAISALWISNL--NQORALARSLIVYQMWAEAVRFRFI 206  
 149 DIERVAGELREELILGMDPLENAISALWISNL--NQORALARSLIVYQMWAEAVRFRFI 206  
 DB 149 DIERVAGELREELILGMDPLENAISALWISNL--NQORALARSLIVYQMWAEAVRFRFI 206

RESULT 4  
 Q94BW5 PRELIMINARY; PRT; 581 AA.  
 AC 094BW5:  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomim I precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).

Query Match	38.8%	Score 395.5	DB 10	Length 549
Best Local Similarity	50.0%	Pred. No. 9.5e-29		
Matches	94	Conservative 27	Mismatches 60	Indels 7
				Gaps 5
DR InterPro: IPR000772; Ricin_B_lectin.				
DR InterPro: IPR001574; RIP.				
DR Pfam: PF00652; Ricin_B_lectin; 5.				
DR Pfam: PF00161; RIP; 1.				
DR PRINTS: PR00396; SHIGARICIN.				
DR SMART: SM00458; RICIN; 2.				
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.				
DR Hydrolyase; Toxin.				
FW NON_TER	1			
FW	1			
SEQUENCE	549 AA;	60648 MM;	02607FEB607CA44B0	CRC64;
Query Match	38.8%	Score 395.5	DB 10	Length 549
Best Local Similarity	50.0%	Pred. No. 9.5e-29		
Matches	94	Conservative 27	Mismatches 60	Indels 7
				Gaps 5
DR InterPro: IPR000772; Ricin_B_lectin.				
DR InterPro: IPR001574; RIP.				
DR Pfam: PF00652; Ricin_B_lectin; 5.				
DR Pfam: PF00161; RIP; 1.				
DR PRINTS: PR00396; SHIGARICIN.				
DR SMART: SM00458; RICIN; 2.				
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.				
DR Hydrolyase; Toxin.				
FW NON_TER	1			
FW	1			
SEQUENCE	549 AA;	60648 MM;	02607FEB607CA44B0	CRC64;

Query Match	34.4%	Score 350.5	DB 10	Length 563
Best Local Similarity	41.6%	Pred. No. 1.8e-24		
Matches	79	Conservative	40	Mismatches 56; Indels 15; Gaps 4
QY	6	YPIINFTTGAIVVOSTNFIKAVVGRRLTGADVHEHPIVLVNNVGLPIKINOFILVELSNH	65	
Db	28	YPSVSNFLAGASATIRDFLKNLRIRVATGTGEVNGVPLRRESEVQVKNRFLVRLTN	87	
QY	66	AEISVTLADVTNATVYVGRAGNSAYFFHPNODAEAT--TILFTDVONRTFAFGNY	123	
Db	88	NGATVTLADVTNATVYVAFSANNASYFF-----KDATQLOKSNLFVQTR-OHTLPTQNY	141	
QY	124	DRLEQAGNIRENIEELGNGPLEEASATYVSTGQTOLPTLANSFICIOISEARFOY	183	
Db	142	DNIEETAGRRRSEIEGLSPDLGATISLYDE-----SVANSLVLIOMSEARFRY	194	
QY	184	IEGEMFRIR	193	
Db	195	IEQEVRSIQ	204	
RESULT 7				
Q9AVR2	ID	PRELIMINARY:	PRT:	564 AA.
AC	Q9AVR2			
AD	Q9AVR2			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, last annotation update)			
DE	Ribosome-inactivating protein precursor (EC 3.2.2.22) (rRNA N-glycosidase).			
GN	EBU1.			
OS	Sambucus ebulus.			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Asteridae; Campanulids; Dipsacales; Adoxaceae; Sambucus.			
OK	NCBI_TaxID=26503;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=leaf;			
RA	Gilbes T., Iglesias R., Perez Y., Ferreras J.M., Cliores L.;			
RT	"Molecular cloning of ebulin 1."			
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
CC	-1- CATALYTIC ACTIVITY: ENDOPHYLISIS OF THE N-GLYCOSIDIC BOND AT ONE SPECIFIC ADENOSINE ON THE 28S RNA.			
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.			
DR	EMBL: AJ400822; CAC33178.1; --			
DR	HSSP: P02879; 2AAT			
DR	InterPro: IPR000772; Ricin_B_lectin.			
DR	InterPro: IPR001574; RIP			
DR	Pfam: PF00652; Ricin_B_lectin; 6.			
DR	Pfam: PF00161; RIP; 1.			
DR	PRINTS: PR00396; SHIGARICIN.			
DR	SMART: SM00458; RICIN; 2.			
DR	PROSITE: PS00231; RICIN_B_LECTIN; 2.			
DR	PROSITE: PS00275; SHIGA_RICIN; 1.			
KW	Glycosidase; Hydrolase; Signal; Toxin.			
FT	SIGNAL	1	25	POTENTIAL.
FT	CHAIN	26	298	EBULIN L A-CHAIN.
FT	CHAIN	299	564	EBULIN L B-CHAIN.
FT	SEQUENCE	564 AA;	62694 MW;	8251681A0DB55C88 CMC64;
Query Match	34.1%	Score 347.5	DB 10	Length 564
Best Local Similarity	41.1%	Pred. No. 3.4e-24		
Matches	78	Conservative	41	Mismatches 56; Indels 15; Gaps 5
QY	6	YPIINFTTGAIVVOSTNFIKAVVGRRLTGADVHEHPIVLVNNVGLPIKINOFILVELSNH	65	
Db	28	YPSVSNFLAGASATIRDFLKNLRIRVATGTGEVNGVPLRRESEVQVKNRFLVRLTN	87	
QY	66	AEISVTLADVTNATVYVGRAGNSAYFFHPNODAEAT--TILFTDVONRTFAFGNY	123	
Db	88	NGATVTLADVTNATVYVAFSANNASYFF-----KDATQLOKSNLFVQTR-OHTLPTQNY	141	
QY	124	DRLEQAGNIRENIEELGNGPLEEASATYVSTGQTOLPTLANSFICIOISEARFOY	183	
Db	142	DNIEETAGRRRSEIEGLSPDLGATISLYDE-----SVANSLVLIOMSEARFRY	194	
QY	184	IEGEMFRIR	193	
Db	195	IEQEVRSIQ	204	

QY	12	DRLEOLAGNLRREIETLGNLELEAISAALYYSGTGCTPLTARSFIICIOIMISARFOY	183
Db	142	DNLEETATAGTRRESIEELGNPLDGLAITSMTY--DGG-----VARSLLVLIQWPPAAAFRY	194
QY	184	IEGEMRTIR	193
Db	195	IEQEVRSIQ	204
RESULT 8			
Q06076			
ID	Q06076	PRELIMINARY;	PRT; 528 AA.
AC	Q06076;		
DT	01-NOV-1996 (TrEMBLrel. 01, Created)		
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Abtin-d (EC 3.2.2.22) (rRNA N-glycosidase) (Frigment).		
OS	Aburys precatotius (Indian licoice) (Crab's eye).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.		
OX	NCBI_TaxID=3816;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=93132798; PubMed=8421313;		
RA	Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.,		
RT	"Primary structure of three distinct isobarnins determined by cDNA		
RT	sequencing: conservation and significance."		
RL	J. Mol. Biol. 229:263-267(1993).		
CC	-I- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE		
CC	SPECIFIC ADENOSINE ON THE 28S RRNA.		
CC	-I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.		
DR	EMBL; M98346; AAA32626.1; -		
DR	HSSP; P11440; IABR.		
DR	InterPro; IPR000772; Ricin_B_lectin.		
DR	InterPro; IPR001574; RIP.		
DR	Pfam; PF00652; Ricin_B_lectin; 6.		
DR	Pfam; PF00161; RIP; 1.		
DR	PRINTS; PR00396; SHIGARICIN.		
DR	SMART; SM00458; RICIN; 2.		
DR	PROSITE; PS50231; RICIN_B_LECTIN; 2.		
DR	PROSITE; PS00275; SHIGA_RICIN; 1.		
KW	Hydrolase; Toxin.		
FT	NON_TER	1	
FT	NON_TER	528	
SO	SEQUENCE	528 AA; 58870 MM; 62ED42B8FBE6F8 CRC64;	
Query Match		33.8%; Score 344.5; DB 10; Length 528;	
Best Local Similarity		45.0%; Pred. No. 5,9e-24;	
Matches	85; Conservative	24; Mismatches	71; Indels 9; Gaps 4
QY	5	QYPLINTTGAAGVQSYNTFIRAVRRLTGADVGHETIPVLPNRGLILNORFLIVELSN	64
Db	1	QDVIKFTTGTGATSSQYKQFTELRLKRLGG--LHNDIPVDPPTTVEERNRYTIVELSN	58
QY	65	HAELSTVLADVYNAVYVCGRAGNSAYFFHPNDEADEATHTLFTDVONRYTFAFGNYD	124
Db	59	SERSISIVGIDVYNAVYVVRAGSSQYEL--RDAPASATYLFPGTQ-RISLRDGSYG	114
QY	125	RLEOLAGNLRREIETLGNLELEAISAALYYSGTGCTPLTARSFIICIOIMISARFOY	184
Db	115	DLEWHAQHTREETSIGLQALTAHS--FLRSGASNDEKARTLLIVIIOMASEARFYCI	177
QY	185	EGEMRTIR	193
Db	172	SNRGVYSIR	180
RESULT 9			
Q04KE4			
ID	Q04KE4	PRELIMINARY;	PRT; 289 AA.
AC	Q04KE4;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		



```

DR 01-DEC-2001 (Tremblrel. 19, last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, last annotation update)
DE Trichosanthis precursor (EC 3.2.2.22) (rRNA N-glycosidase).
GN TCS.
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RA Yuan H., Wang Y., An C., Chen Z.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDONUCLEOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL: AF367252; AAK52960.1; -.
DR InterPro: IPR001574; RFP.
DR Pfam: PF00161; RFP.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 23
FT CHAIN 24 270
FT SEQUENCE 289 AA; 31706 MW; A6D5602549CA5657 CRC64;

Query Match 33.1%; Score 340; DB 10; Length 289;
Best Local Similarity 39.5%; Pred. No. 6.8e-24;
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;

QY 9 INFTTGATVOSTYNTFIRAVRGLTGADVREHPIVLPNRYGPIPNORFLVELSNHAEI 68
DB 25 VSRRLSGATSSSYGVFISNLKALPNERKL-YDIPLL--RSLPGSORVALHILTNVYADE 81
QY 69 SVTLALDVTNAVYVYGRAGNSAYFFHPDNOEDA-EAITHLEFTDVONRYTFAGGNYDRLE 127
DB 82 TISVAIDVTNVYIMGRAGDTSYFF--NEASATEAKYFKRSMKRTLPYSGNTERLQ 138
QY 128 QLAGNLRNIEIENGNGLEPAISALYYSTGCTQLPLARSFICIQMISEARFQYIEGE 187
DB 139 TAAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEARXKFIQEQ 193
QY 188 MRTRI 192
DB 194 IGRKV 198

RESULT 10
Q9LRE3 PRELIMINARY; PRT; 247 AA.
AC Q9LRE3;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DE 01-OCT-2002 (Tremblrel. 22, last annotation update)
DE Trichobakin (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).
GN TKB.
OS Trichosanthes sp. Bac Kan 8-98.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=118182;
RN [1]
RP SEQUENCE FROM N.A.
RA Phan Van Chi, Nguyen Thuy Ha, Le Tran Binh;
RT "Genomic DNA Clone for mature typ-1 ribosome-inactivating protein from
RT Trichosanthes sp. sample 01 Bac Kan 8-98 Vien CSMH (Hanoi).";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL: AA039324; BAA9230.1; -.
DR HSP: P09989; IMRJ.
DR InterPro: IPR001574; RFP.

```

```

DR Pfam: PF00161; RFP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Toxin.
FT NON_TER 1 1
FT NON_TER 247 247
FT SEQUENCE 247 AA; 27199 MW; 89811AC32892F03F CRC64;

Query Match 33.1%; Score 338; DB 10; Length 247;
Best Local Similarity 40.0%; Pred. No. 8.5e-24;
Matches 74; Conservative 46; Mismatches 53; Indels 12; Gaps 5;

QY 9 INFTTGATVOSTYNTFIRAVRGLTGADVREHPIVLPNRYGPIPNORFLVELSNHAEI 68
DB 2 VSRRLSGATSSSYGVFISNLKALPNERKL-YDIPLL--RSLPGSORVALHILTNVYADE 58
QY 69 SVTLALDVTNAVYVYGRAGNSAYFFHPDNOEDA-EAITHLEFTDVONRYTFAGGNYDRLE 127
DB 59 TISVAIDVTNVYIMGRAGDTSYFF--NEASATEAKYFKRSMKRTLPYSGNTERLQ 115
QY 128 QLAGNLRNIEIENGNGLEPAISALYYSTGCTQLPLARSFICIQMISEARFQYIEGE 187
DB 116 TAAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEARXKFIQEQ 170
QY 188 MRTRI 192
DB 171 IGRKV 175

RESULT 11
Q41216 PRELIMINARY; PRT; 289 AA.
AC Q41216;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DE 01-OCT-2002 (Tremblrel. 22, last annotation update)
DE Trichosanthis (EC 3.2.2.22) (rRNA N-glycosidase).
GN TRICHOSANTHIN, TCS.
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94271613; PubMed=8003348;
RA Zheng H., Wang B., Shaw P., Yeung H.;
RT "[Cloning and DNA sequencing of the gene encoding trichosanthis].";
RL I Chuan Hsueh Pao 21:42-51(1994).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL: S70176; AAB31048.1; -.
DR HSP: P09989; IMRJ.
DR InterPro: IPR001574; RFP.
DR Pfam: PF00161; RFP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Toxin.
FT SEQUENCE 289 AA; 31650 MW; 286AC14D48BCA175 CRC64;

Query Match 33.1%; Score 338; DB 10; Length 289;
Best Local Similarity 39.5%; Pred. No. 1.1e-23;
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;

QY 9 INFTTGATVOSTYNTFIRAVRGLTGADVREHPIVLPNRYGPIPNORFLVELSNHAEI 68
DB 25 VSRRLSGATSSSYGVFISNLKALPNERKL-YDIPLL--RSLPGSORVALHILTNVYADE 81
QY 69 SVTLALDVTNAVYVYGRAGNSAYFFHPDNOEDA-EAITHLEFTDVONRYTFAGGNYDRLE 127
DB 82 TISVAIDVTNVYIMGRAGDTSYFF--NEASATEAKYFKRSMKRTLPYSGNTERLQ 138

```

QY 128 QLAGNRENIENGPLEAISALYYSTGCTOLPLANSFTICOMISEARFOYTEGE 187  
 Db 139 TAAGKIRENIPGLPALDSAITPLFYNNAN-----SAASALMWLIOSTSEARFYKTEEOO 193  
 QY 188 MRRIR 192  
 Db 194 IGRV 198

## RESULT 12

Q38760 PRELIMINARY; PRT; 252 AA.

AC Q38760; 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE ABRIN-E (RRNA N-glycosidase) (EC 3.2.2.22) (Fragment).  
 GN RIP.

OS Abrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abraceae; Abrus.  
 OC NCBI\_TaxID=3816;  
 RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LEAF;  
 MEDLINE=91201329; PubMed=2016300;

RA Evensen G., Mathiesen A., Sundan A.;  
 RT "Direct molecular cloning and expression of two distinct abrin A-  
 chains";

RL J. Biol. Chem. 266:6848-6852(1991).  
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.

CC ABRIN-A IS MORE TOXIC THAN RICIN.  
 CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
 FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
 PRECEDES ENOCYTOSIS.

CC CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.

CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -1- SIMILARITY: TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING  
 PROTEINS. BELONGS TO TYPE 2 RIP.

DR EMBL: X54872; CA38654.1; -;  
 DR HSSP: P11140; IABR.  
 DR InterPro: IPR001574; RIP.

DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.

FT Hydrolase; Glycosidase; Toxin; Repeat; Glycoprotein; Lectin.  
 FT CHAIN 1 252 ABRIN E, A CHAIN (BY SIMILARITY).  
 FT NON\_TER 252  
 SQ SEQUENCE 252 AA; 28309 MW; BFFC846B9E92B5DE CRC64;

Query Match 33.0%; Score 336.5; DB 10; Length 252;  
 Best Local Similarity 44.9%; Pred. No. 1.2e-23;  
 Matches 83; Conservative 24; Mismatches 69; Indels 9; Gaps 4;

QY 9 INFTTAGATVOSTYNTFIRAVRGLTTGADVRHEIPVLPNRVGLPINORFIVELSNHAE 68  
 Db 6 IKFSTEGATVOSTYNTFIRAVRGLTTGADVRHEIPVLPNRVGLPINORFIVELSNHAE 63  
 QY 69 SVTLALDVTNAYVVGIRAGNSAYFFHPDQDAEALFTFTDVQNRRTFAFGAGYDRLEQ 128  
 Db 64 SIEVGIDVTNAYVVGIRAGNSAYFFHPDQDAEALFTFTDVQNRRTFAFGAGYDRLEQ 119  
 QY 129 LAGNIRENIEGNGPLEAISALYYSTGCTOLPLANSFTICOMISEARFOYTEGE 188  
 Db 120 MAHQTRREGISGLQALTHAIS---FLRSGASNDKAKATLLVLIQMSSEARFYISNRV 176  
 QY 189 RTRIR 193

Db 177 GVSIR 181

## RESULT 13

Q94552 PRELIMINARY; PRT; 563 AA.

AC Q94552; 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA  
 DE N-glycosidase).  
 GN AVL.

OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OC NCBI\_TaxID=4202;  
 RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Leaf;

RA Van Damme E.J.M.;  
 RT "Characterization and cloning of lectins and ribosome-inactivating  
 proteins from Sambucus nigra leaves";

RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RNA.

CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AF409135; AAL04123.1; -;  
 DR InterPro: IPR007772; Ricin\_B\_lectin.

DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR PRINTS: PR00396; SHIGARICIN.

DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 DR Hydrolase; Toxin.

SQ SEQUENCE 563 AA; 62242 MW; 07F7CBDC33BF10 CRC64;

Query Match 32.4%; Score 330.5; DB 10; Length 563;  
 Best Local Similarity 39.5%; Pred. No. 1.4e-22;  
 Matches 75; Conservative 39; Mismatches 61; Indels 15; Gaps 4;

QY 6 YPIINFTTAGATVOSTYNTFIRAVRGLTTGADVRHEIPVLPNRVGLPINORFIVELSNH 65  
 Db 28 YPSVSNINDGAKSATYRDELNLNMRKTVANGTEVNGDLPVLRRESVYKSRFVLVPLTNY 87  
 QY 66 AELSVTLALDVTNAYVVGIRAGNSAYFFHPDQDAEAL--THEFTDVQNRRTFAFGAGNY 123  
 Db 88 NGWTVTLADVTNAYVVGIRAGNSAYFFHPDQDAEAL--KATEVOKSNLTVGTRKN--TISFTGNY 141  
 QY 124 DRLQLAGNIRENIEGNGPLEAISALYYSTGCTOLPLANSFTICOMISEARFOYTEGE 183  
 Db 142 DNLETAANRRRESIEGSPFDGATITSLYHGD-----SVAKSLVLIQMSSEARFY 194  
 QY 184 IEGEMRTIR 193  
 Db 195 IEGEVRRSLQ 204

## RESULT 14

Q8GT32 PRELIMINARY; PRT; 563 AA.

AC Q8GT32;

DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

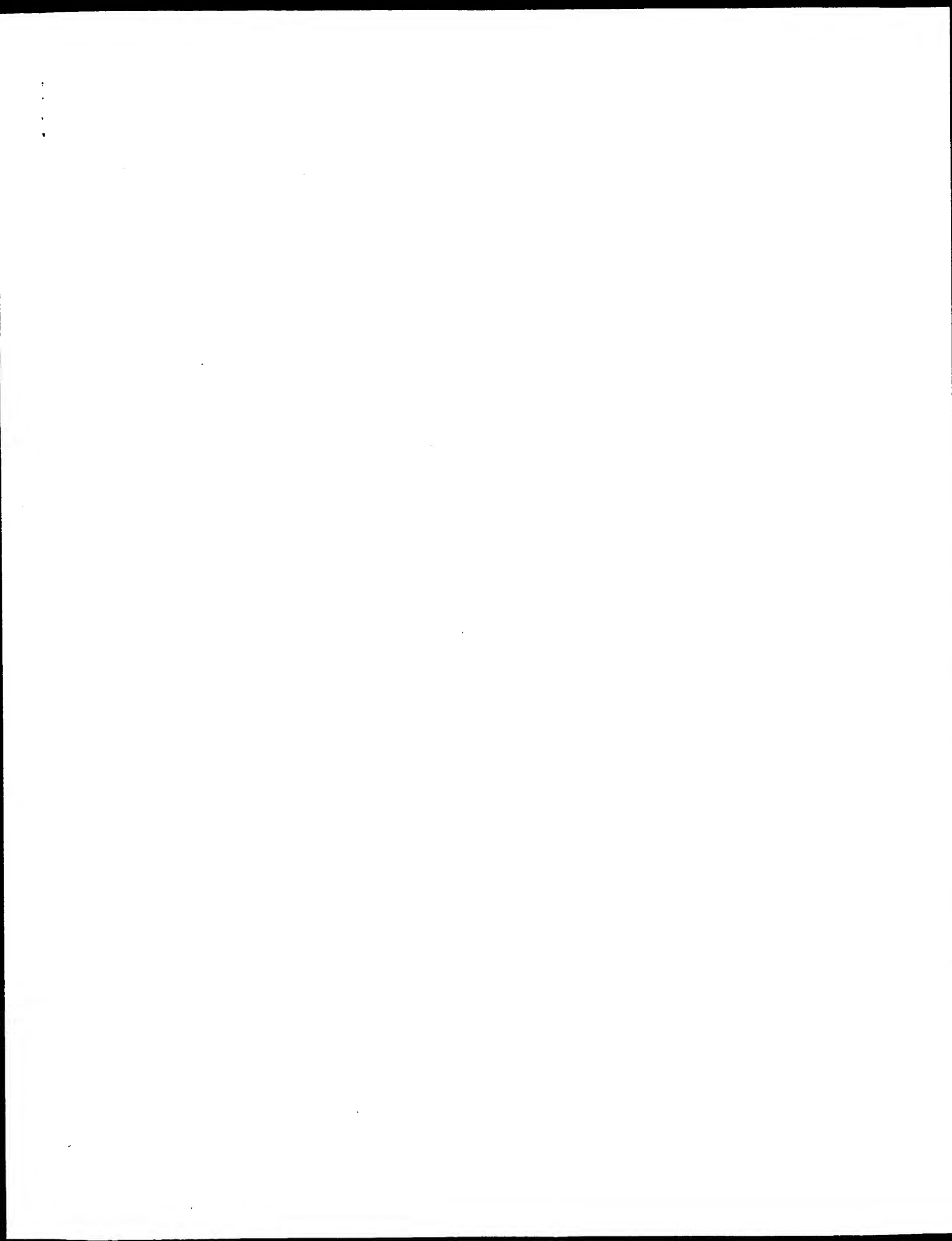
DE Type 2 ribosome-inactivating protein nigrin I precursor  
 DE (EC 3.2.2.22).  
 OS Sambucus nigra (European elder).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta: Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RA Girbes T., Arias F.J., Antolin P.;  
 RT "Characterization and molecular cloning of Nigrlin 1, a type two  
 RT ribosome-inactivating protein from leaves of elder (Sambucus nigra).";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF249280; AAB6130.1; -  
 KM Hydrolyase; Glycosidase.  
 SQ SEQUENCE 563 AA; 62173 MW; 0EB236421FC5E04F CRC64;  
 Query Match 32.4%; Score 330.5; DB 10; Length 563;  
 Best Local Similarity 39.5%; Pred. No. 1.4e-22;  
 Matches 75; Conservative 39; Mismatches 61; Indels 15; Gaps 4;  
 QY 6 YPIINFTAGATVQSYTNFIRAVRGLTTGADVRIHEIPVLPNVRGLPINORFIVELSNH 65  
 DB 28 YPSVSEFMDGAKSATYEDFLSNLRKTYATCTYEYNGPLVLRSEVQVSRFLVPLTNV 87  
 QY 66 AELSVTLALDVTNAVYVGYAGNSAYFEFHEDNOEDAEAT--TLFTDVQNRITPAFGN 123  
 DB 88 NGNTVTLADVTLNLYVAEFGNANSYFE---KDATVQKSNLFEVGTQKQ--TLSTGNV 141  
 QY 124 DRLEQLAGNLRNIELGNGPLEEAIASALYYSTGTQPLTARSFICITOMISEARPOY 183  
 DB 142 DNELETANTRRESIELGSPSLDCAITSLYHGD-----SVARSLVLIOMVSEARFR 194  
 QY 184 IEEMRTIR 193  
 DB 195 IEQEVRRSLQ 204  
 RESULT 15  
 ID 038761 PRELIMINARY; PRT; 252 AA.  
 AC 038761; 096234;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Ricin A-chain type 73 (EC 3.2.2.22) (rRNA N-glycosidase)  
 DE (Fragment).  
 GN RLP.  
 OS Atrius precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Atrius.  
 OX NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RA Evensen G., Mathiesen A., Sundan A.;  
 RL Submitted (OCT-1990) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 2-252 FROM N.A.  
 RC TISSUE=Seed;  
 RX MEDLINE=94139756; PubMed=8307038;  
 RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;  
 RT "Cloning and expression of three abrin A-chains and their mutants  
 RT derived by site-specific mutagenesis in Escherichia coli.";  
 RL Eur. J. Biochem. 219:83-87(1994).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC EMBL: X54873; CAA38655.1; -  
 DR EMBL: X76720; CAA54138.1; -  
 DR HSSP: P11140; IABR.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP, 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.

KM Glycosidase; Hydrolase; Toxin.  
 FT NON\_TER 252 252  
 SQ SEQUENCE 252 AA; 28229 MW; 187B8B4E134ACE5 CRC64;  
 Query Match 32.3%; Score 329.5; DB 10; Length 252;  
 Best Local Similarity 42.8%; Pred. No. 5.5e-23;  
 Matches 80; Conservative 28; Mismatches 66; Indels 13; Gaps 5;  
 QY 9 INFTAGATVQSYTNFIRAVRGLTTGADVRIHEIPVLPNVRGLPINORFIVELSNH 68  
 DB 6 IKFSTEGATQSYKQFTEALRRLRG--LIHDIVLPDPTTLOERNRRTIVELSNSTE 63  
 QY 69 SVTLADVTNAVYVGYAGNSAYFEFH--PDNOEDAEATLFTDVQNRITPAFGN 126  
 DB 64 SIEVGIDVTNAVYVGYAGNSAYFEFH--PDNOEDAEATLFTDVQNRITPAFGN 117  
 QY 127 EQLAGNLRNIELGNGPLEEAIASALYYSTGTQPLTARSFICITOMISEARPOY 186  
 DB 118 ERMHQSROQIPLGLQALTHGIS---FFRSGNDNEKARTLITVLIOMVSEARFR 174  
 QY 187 EMRTIR 193  
 DB 175 RVRSIQ 181

Search completed: September 16, 2003, 11:49:14  
 Job time : 31.5828 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:29:47 ; Search time 31.3059 Seconds  
(without alignments)  
953.195 Million cell updates/sec

Title: US-10-083-336a-4

Perfect score: 966  
Sequence: 1 IFPKQYPIINFTTAAATVQS.....ARFYIEGEMRTIRYNRRS 188

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: A\_Geneseq\_19Jun03:\*

2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*

3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*

4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*

5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*

6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*

7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*

8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*

9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*

10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*

11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*

12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*

13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*

14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*

15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*

16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*

17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*

18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*

19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*

20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*

21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*

22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*

23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*

24: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	951	98.4	267	14	AA1980.DAT:*
2	951	98.4	267	16	AA1981.DAT:*
3	951	98.4	267	18	AA1982.DAT:*
4	951	98.4	267	20	AA1983.DAT:*
5	951	98.4	267	22	AA1984.DAT:*
6	951	98.4	267	24	AA1985.DAT:*
7	951	98.4	267	26	AA1986.DAT:*
8	951	98.4	267	28	AA1987.DAT:*
9	951	98.4	267	30	AA1988.DAT:*

#### ALIGNMENTS

10	951	98.4	565	6	AA1980.DAT:*	Sequence of prepro
11	951	98.4	565	22	AA1981.DAT:*	Castor bean prepro
12	951	98.4	565	22	AA1982.DAT:*	Modified castor be
13	951	98.4	576	8	AA1983.DAT:*	Sequence of Ricinu
14	951	98.4	576	18	AA1984.DAT:*	Castor bean ricin
15	951	98.4	576	18	AA1985.DAT:*	Castor bean ricin
16	951	98.4	576	21	AA1986.DAT:*	Ricinus communis r
17	951	98.4	576	22	AA1987.DAT:*	Castor bean prepro
18	951	98.4	576	22	AA1988.DAT:*	Castor bean prepro
19	947	98.0	565	9	AA1989.DAT:*	Castor bean prepro
20	946	97.9	565	7	AA1990.DAT:*	Biosynthetic multi
21	944	97.7	565	14	AA1991.DAT:*	Preprolin. Ricl
22	944	97.7	574	14	AA1992.DAT:*	Sequence of ricin-
23	942	97.5	574	10	AA1993.DAT:*	Sequence of Ricinu
24	941	97.4	567	13	AA1994.DAT:*	DNA sequence of r1
25	941	97.4	567	14	AA1995.DAT:*	Ricin A from pR1C1
26	941	97.4	534	14	AA1996.DAT:*	Amino acid sequenc
27	940	97.3	332	11	AA1997.DAT:*	Sequence of G-RT.
28	934	96.7	267	16	AA1998.DAT:*	Ricin A gene produ
29	918	95.0	267	14	AA1999.DAT:*	Ricin A chain (RTA
30	861.5	89.2	540	18	AA2000.DAT:*	Ricin A. Syntheti
31	861.5	89.2	540	18	AA2001.DAT:*	Castor oil plant a
32	801.5	83.0	534	8	AA2002.DAT:*	R. communis agglut
33	718	74.3	280	10	AA2003.DAT:*	Sequence of Ricin
34	336	34.8	247	16	AA2004.DAT:*	Ricin agglutinin A
35	336	34.8	247	16	AA2005.DAT:*	Trichosanthin anti
36	336	34.8	248	11	AA2006.DAT:*	Amino acid sequenc
37	336	34.8	248	13	AA2007.DAT:*	Synthetic alpha-tr
38	336	34.8	267	18	AA2008.DAT:*	Mature alpha-trich
39	336	34.8	267	18	AA2009.DAT:*	Trichosanthin (a r
40	336	34.8	289	11	AA2010.DAT:*	Trichosanthin. Tr
41	336	34.8	289	13	AA2011.DAT:*	Trichosanthin from
42	336	34.8	289	13	AA2012.DAT:*	Trichosanthin prot
43	336	34.8	289	14	AA2013.DAT:*	Encodes chinese cu
44	336	34.8	289	15	AA2014.DAT:*	Alpha-trichosanthl
45	336	34.8	289	18	AA2015.DAT:*	Chinese cucumber a

RESULT 1  
ID AAR37290 standard; protein; 267 AA.

AC AAR37290:

DT 25-MAR-2003 (updated)  
DT 09-JAN-2003 (updated)  
DT 13-SEP-1993 (first entry)

DE Ricin A chain.

KM Type II ribosome-inactivating protein; type II RIP; gelonin;  
KW monomordin; immunconjugate; autoimmune disease; cell killing; toxin.

OS Unidentified.

PN WO309130-A1.

PD 13-MAY-1993.

PF 04-NOV-1992; 92WO-US09487.

PR 04-NOV-1991; 91US-0787567.

PR 19-JUN-1992; 92US-0901707.

PA (XOMA ) XOMA CORP.

PI Bernhard SL., Better MD, Carroll SF, Lane JA, Lei SP;

DR WPL: 1993-167617/20.

PT Analogues of type I ribosome inactivating protein - useful as  
 PT cytotoxic agents, immuno toxins for treating autoimmune diseases,  
 PT cancer, graft versus host disease and selective cell killing in-vivo  
 PS Claim 1: Page 92: 163pp: English.

XX The invention covers analogues of Type I RIPS. Ricin is a Type II  
 CC RIP whose A chain is homologous to plant type I RIPS. The analogues  
 CC of the invention have a cysteine available for intermolecular  
 CC disulphide bonding at an amino acid position corresp. to a position  
 CC not naturally available for bonding; the cys residue is located in  
 CC the C-terminal region of the analogue between a position corresp. to  
 CC amino acid 251 and the C-terminus of ricin A chain. The analogues are  
 CC pref. joined via a disulphide linkage to a molecule which specifically  
 CC binds to a target cell, e.g. an antibody fragment.  
 CC (updated on 09-JAN-2003 to add missing OS field.)  
 CC (updated on 25-MAR-2003 to correct PN field.)

XX Sequence 267 AA;

Query Match 98.4%; Score 951; DB 14; Length 267;  
 Best Local Similarity 94.9%; Pred. No. 9.2e-94;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFPKQYPIINFTTAGATVQSTNFIKAVRGRLT-----VLPNRVGLPINORFLLV 50  
 DB 1 IFPKQYPIINFTTAGATVQSTNFIKAVRGRLTGGADVHEIPVLPNRVGLPINORFLLV 60  
 QY 51 ELSNHAELSVTLADVTNNAVYVGRAGNSAYFFHPDNOEDAEATHTLFTDVQNRVTFAG 110  
 DB 61 ELSNHAELSVTLADVTNNAVYVGRAGNSAYFFHPDNOEDAEATHTLFTDVQNRVTFAG 120  
 QY 111 GNYDRLEQLAGNLRNENIELGNGPLEEASALYYSTGCTOLPTLARSFFICIONMISEAR 170  
 DB 121 GNYDRLEQLAGNLRNENIELGNGPLEEASALYYSTGCTOLPTLARSFFICIONMISEAR 180  
 QY 171 FOYIEGEMRTIRIRNRRS 188  
 DB 181 FOYIEGEMRTIRIRNRRS 198

RESULT 2  
 AAR63902 standard; protein: 267 AA.

XX AAR63902;  
 AC  
 XX 25-MAR-2003 (updated)  
 DT 27-JUL-1995 (first entry)  
 DE Ricin A-chain (RTA).  
 XX  
 KW Ricin A chain; RTA; ribosome-inactivating proteins; RIPS;  
 KW cytotoxic therapeutic agents; autoimmune disease; cancer;  
 KW graft-versus-host disease.  
 XX  
 OS Ricinus communis.  
 XX  
 PN MO9426910-A1.  
 PD 24-NOV-1994.  
 PF 12-MAY-1994; 94WO-US05348.  
 PR 12-MAY-1993; 93US-0064691.  
 PA (XOMA ) XOMA CORP.  
 PI Better MD, Carroll SS, Studnicka GM, Carroll SF;  
 DR WPI; 1995-006804/01.  
 XX  
 PT Polynucleotide(s) encoding type I ribosome-inactivating proteins

PT - which are suitable for use as components of cytotoxic  
 PT therapeutic agents.

PS Example 3; Fig 1; 221pp: English.

XX AAR63902 is the ricin A chain gene product, it is analogous to the  
 CC ribosome-inactivating proteins (RIPs) described in AAR63903-R63911.  
 CC RIPs are the key components of cytotoxic therapeutic agents (CTAs),  
 CC which include gene fusion products and immunconjugates. CTAs may  
 CC be used to selectively eliminate any cell type to which a RIP  
 CC component is targeted, by the specific binding capacity of the  
 CC second component of the agent. They can be used in the treatment  
 CC of diseases where the elimination of a particular cell type is  
 CC desired, such as autoimmune disease, cancer and graft-versus-host  
 CC disease.  
 CC (updated on 25-MAR-2003 to correct PN field.)

XX Sequence 267 AA;

Query Match 98.4%; Score 951; DB 16; Length 267;  
 Best Local Similarity 94.9%; Pred. No. 9.2e-94;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFPKQYPIINFTTAGATVQSTNFIKAVRGRLT-----VLPNRVGLPINORFLLV 50  
 DB 1 IFPKQYPIINFTTAGATVQSTNFIKAVRGRLTGGADVHEIPVLPNRVGLPINORFLLV 60  
 QY 51 ELSNHAELSVTLADVTNNAVYVGRAGNSAYFFHPDNOEDAEATHTLFTDVQNRVTFAG 110  
 DB 61 ELSNHAELSVTLADVTNNAVYVGRAGNSAYFFHPDNOEDAEATHTLFTDVQNRVTFAG 120  
 QY 111 GNYDRLEQLAGNLRNENIELGNGPLEEASALYYSTGCTOLPTLARSFFICIONMISEAR 170  
 DB 121 GNYDRLEQLAGNLRNENIELGNGPLEEASALYYSTGCTOLPTLARSFFICIONMISEAR 180  
 QY 171 FOYIEGEMRTIRIRNRRS 188  
 DB 181 FOYIEGEMRTIRIRNRRS 198

RESULT 3  
 AAM25136 standard; protein: 290 AA.

XX AAM25136;  
 AC  
 XX 25-MAR-2003 (updated)  
 DT 02-DEC-1997 (first entry)  
 DE Ricin A-chain ribosome inhibitory protein inactive precursor.  
 XX  
 KW Maize; proRIP; ribosome inactivating protein; alpha; beta subunit;  
 KW internal linker; Barley translation inhibitor; Trichosanin;  
 KW Ricin A-chain; Abrin-A-A-chain; Saporin; SLT-1; Luifin A; MAP;  
 KW Ricinus communis agglutinin; Momordin; PAP-S; Luifin-B; Dianthin 30;  
 KW therapeutic toxin; tumor cell targeted; protein synthesis inhibitor;  
 KW post-translational modification; cancer; neoplasia; HIV; AIDS;  
 KW human immunodeficiency virus; acquired immune deficiency syndrome.  
 XX  
 OS Synthetic.  
 XX  
 PN US5646026-A.  
 PD 08-JUL-1997.  
 PF 07-JUN-1995; 95US-0485286.  
 PR 09-DEC-1992; 92US-0987927.  
 PR 11-JUN-1990; 90US-0535636.  
 PR 26-JAN-1995; 95US-0378761.  
 PR 07-JUN-1995; 95US-0485286.  
 XX  
 PA (DOWC ) DOWELANCO.

XX Hey TD, Morgan AER, Walsh TA;  
 PI WPI: 1997-362934/33.  
 XX  
 DR  
 XX  
 PT DNA encoding pro-ribosome inactivating proteins - inactive  
 PT precursors of ribosome inactivating proteins; can be expressed in  
 PT eukaryotic cells without causing cell death  
 XX  
 PS Claim 4; Column 91-94; 186pp; English.  
 XX  
 CC AAM25136 represents a Ricin A-chain ribosome inhibitory protein (RIP)  
 CC which was engineered to contain a selectively removable internal peptide  
 CC linker sequence separating the alpha and beta units of the RIP. When  
 CC separated the two units regain activity and are capable of inactivating  
 CC eukaryotic ribosomes and hence preventing protein production. Many  
 CC different RIPs may be produced with an internal linker including  
 CC maize RIP, Trichosanthin, Ricin A-chain, Abrin-A chain and  
 CC Saporin. The RIPs can be used in the construction of therapeutic  
 CC toxins targeted to specific cells such as tumour cells via the  
 CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.  
 CC A further use is in HIV therapy (see US4869903). There is interest  
 CC in expressing RIP recombinantly in host eukaryotic cells, because of  
 CC the capacity to provide correct post-translational processing. However,  
 CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting  
 CC in cell death. Since the inactive RIP proteins are not cytotoxic to  
 CC eukaryotic cells, they can be recombinantly expressed in such cells and  
 CC then converted to active RIP proteins.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC  
 XX  
 SQ Sequence 290 AA;  
 Query Match 98.4%; Score 951; DB 18; Length 290;  
 Best Local Similarity 94.9%; Pred. No. 1e-93;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 1 IFPKOYPIINFTTGAATVOSTTNTFIRAVRGRLT-----VLPNRVGLPINORFTLV 50  
 DB 25 IFPKOYPIINFTTGAATVOSTTNTFIRAVRGRLTGDADVREHIEIPVLPNRVGLPINORFTLV 84  
 QY 51 ELSNHAELSVTLADVTNNAVYVGRAGNSAVFFHPDNOEDAEATHTLFTVONRYTPAFG 110  
 DB 85 ELSNHAELSVTLADVTNNAVYVGRAGNSAVFFHPDNOEDAEATHTLFTVONRYTPAFG 144  
 QY 111 GNYDRLEQLAGNLRENIELGNGPLEEASISALYYSTGTOPLTARSFTICQIMISEAR 170  
 DB 145 GNYDRLEQLAGNLRENIELGNGPLEEASISALYYSTGTOPLTARSFTICQIMISEAR 204  
 QY 171 FOYIEGEMRTIRIRYNRS 188  
 DB 205 FOYIEGEMRTIRIRYNRS 222  
 RESULT 4  
 AAM21699  
 ID AAM21699 standard; Protein: 290 AA.  
 AC AAM21699;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 26-SEP-1997 (first entry)  
 XX  
 DE Ricin A-chain RIP.  
 XX  
 KW pro-ribosome inactivating Protein; proRIP; peptide linker; cancer;  
 KW inactivation; eukaryotic ribosome; alpha fragment; beta fragment;  
 KW inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;  
 KW rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.  
 OS  
 XX  
 XX Ricinus communis.  
 FH Key Location/Qualifiers  
 FT Region 152..162

FTT /note= "Position of possible insertion of internal  
 XX peptide linker sequence"  
 XX  
 XX US5635384-A.  
 XX  
 PD 03-JUN-1997.  
 XX  
 XX  
 PF 26-JAN-1995; 95US-0378761.  
 XX  
 XX  
 PR 09-DEC-1992; 92US-0987927.  
 PR 11-JUN-1990; 90US-0535636.  
 PR 26-JAN-1995; 95US-0378761.  
 XX  
 PA (DOWC ) DOWELANCO.  
 XX  
 PI Hey TD, Morgan AER, Walsh TA;  
 XX  
 DR WPI: 1997-309831/28.  
 XX  
 PT Inactive precursor of maize ribosome-inactivating protein - also  
 PT chimeric ribosome-inactivating protein precursors containing  
 PT internal linker sequences  
 XX  
 PS Claim 2; Column 91-94; 121pp; English.  
 XX  
 CC The sequences given in AAM21698-710 represent Ribosome Inactivating  
 CC Proteins (RIP's), which may be used in the construction of the  
 CC proRIP of the invention. The proRIP has a selectively removable,  
 CC internal peptide linker. The precursor sequence is incapable of  
 CC inactivating eukaryotic ribosomes, but can be converted by removal  
 CC of the linker into a protein having alpha and beta fragments and being  
 CC capable of inactivating eukaryotic ribosomes. RIPs are potent  
 CC inhibitors of eukaryotic protein synthesis. They possess a highly  
 CC specific N-glycosidase activity which cleaves the glycosidic bond of  
 CC adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit  
 CC cellular proliferation of cells, e.g. cancer cells and HIV-infected T  
 CC cells. The inactive proRIP proteins make it possible to provide protein  
 CC synthesis inhibitors with uses in practical and improved ways not before  
 CC possible. The RIP can be used to make cytotoxic conjugates.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC  
 XX  
 SQ Sequence 290 AA;  
 Query Match 98.4%; Score 951; DB 18; Length 290;  
 Best Local Similarity 94.9%; Pred. No. 1e-93;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 1 IFPKOYPIINFTTGAATVOSTTNTFIRAVRGRLT-----VLPNRVGLPINORFTLV 50  
 DB 25 IFPKOYPIINFTTGAATVOSTTNTFIRAVRGRLTGDADVREHIEIPVLPNRVGLPINORFTLV 84  
 QY 51 ELSNHAELSVTLADVTNNAVYVGRAGNSAVFFHPDNOEDAEATHTLFTVONRYTPAFG 110  
 DB 85 ELSNHAELSVTLADVTNNAVYVGRAGNSAVFFHPDNOEDAEATHTLFTVONRYTPAFG 144  
 QY 111 GNYDRLEQLAGNLRENIELGNGPLEEASISALYYSTGTOPLTARSFTICQIMISEAR 170  
 DB 145 GNYDRLEQLAGNLRENIELGNGPLEEASISALYYSTGTOPLTARSFTICQIMISEAR 204  
 QY 171 FOYIEGEMRTIRIRYNRS 188  
 DB 205 FOYIEGEMRTIRIRYNRS 222  
 RESULT 5  
 AAP70097  
 ID AAP70097 standard; protein: 332 AA.  
 AC AAP70097;  
 XX  
 DT 09-APR-1991 (first entry)  
 XX  
 DE Ricin A.

```

XX XX Ricin A; Met-aminopeptidase.
XX XX Escherichia coli.
XX PN EP219237-A.
XX PD 22-APR-1987.
XX PF 19-SEP-1986; 86EP-0307242.
XX PR 06-MAY-1986; 86US-0860330.
XX PR 20-SEP-1985; 85US-0778414.
XX PA (CETU ) CETUS CORP.
XX PI Benbasat A, Bauer KA, Chang S, Chang SY;
XX DR WPI; 1987-110172/16.
XX DR N-PSDB; AAN70152.
XX PT N-terminal methionine free proteins prodn. - by using host
XX PT transformed with vector to express a methionine- amino-peptidase
XX PS Disclosure; Fig. 4; 20pp; English.
XX CC Ricin A may be produced in a form which lacks an N-terminal Met
XX CC using Met-aminopeptidase from E.coli.
XX SO Sequence 332 AA;

Query Match 98.4%; Score 951; DB 8; Length 332;
Best Local Similarity 94.9%; Pred. No. 1.2e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLT-----VLPNRVGLPINQRFILV 50
DB 36 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTGADVRHRIPLVLPNRVGLPINQRFILV 95
QY 51 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAETHLFTDVQNRYPFAFG 110
DB 96 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAETHLFTDVQNRYPFAFG 155
QY 111 GNYDRLEQLAGNLRENIELGNGPLEEASIALYYSTGCTQLPTLARSFTICIMISEAR 170
DB 156 GNYDRLEQLAGNLRENIELGNGPLEEASIALYYSTGCTQLPTLARSFTICIMISEAR 215
QY 171 FOYIEGEMRTIRIRNRRS 188
DB 216 FOYIEGEMRTIRIRNRRS 233

RESULT 6
AAP70838
ID AAP70838 standard; protein; 332 AA.
XX AC AAP70838;
XX DT 25-MAR-2003 (updated)
XX DT 18-FEB-1991 (first entry)
XX DE Sequence of Ricinus communis castor beans ricin toxin (RT or ricin)
XX DE A protein encoded by pR1A123.
XX KW Lectin; toxin protein; cytotoxic; cytostatic; castor bean;
XX KW plant toxin.
XX OS Ricinus communis.
XX FH Key
XX FH Location/Qualifiers
XX FT 1..32
XX FT Region
XX FT /note="Leader"
XX FT 33..302

```

```

FT FT /note="A-chain"
FT FT 315..332
FT FT /note="B-chain"
XX PN EP237676-A.
XX PD 23-SEP-1987.
XX PF 13-NOV-1986; 86EP-0308877.
XX PR 07-MAR-1986; 86US-0837583.
XX PA (CETU ) CETUS CORP.
XX PA (CHIR ) CHIRON CORP.
XX PI Platak M;
XX DR WPI; 1987-265177/38.
XX DR N-PSDB; AAN70519.
XX PT New non-glycosylated ricin precursor and toxin etc. - are prepd.
XX PT by recombinant DNA procedures with specific isolation steps for
XX PT purer and soluble prods.
XX PS Disclosure; Fig 1; 112pp; English.
XX CC The full-length sequences encoding ricin A (AAN70520), ricin D
XX CC (AAN70525), putative ricin E (AAN70526) and RCA (AAN70524) in precursor
XX CC form were obtd. using messenger RNA to obtain a cDNA library, and
XX CC then probing the library to retrieve the desired cDNA inserts. The
XX CC library was probed using the 35-mer given in AAN70514, figure 4 (see
XX CC AAN70520, AAN70521, AAN70522) shows the nucleotide sequences of three
XX CC plasmids contg. cDNA inserts obtd. by probing a cDNA library for
XX CC sequences encoding ricin B using the probe in AAN70517. The cDNA
XX CC inserts can be placed into expression vectors. Site-directed
XX CC mutagenesis may be used to place an ATG start codon and a HindIII
XX CC site at the beginning of the mature protein (see AAN70518). The
XX CC coding sequences of the inserts can be ligated into expression
XX CC vectors contg. the phoA promoter-operator and leader sequence
XX CC (AAN70523) and suitable retroregulators.
XX CC (updated on 25-MAR-2003 to correct PA field.)
XX SO Sequence 332 AA;

Query Match 98.4%; Score 951; DB 8; Length 332;
Best Local Similarity 94.9%; Pred. No. 1.2e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLT-----VLPNRVGLPINQRFILV 50
DB 36 IFFKQYPIINFTTAGATVOSTYNTFIRAVRGRLTGADVRHRIPLVLPNRVGLPINQRFILV 95
QY 51 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAETHLFTDVQNRYPFAFG 110
DB 96 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAETHLFTDVQNRYPFAFG 155
QY 111 GNYDRLEQLAGNLRENIELGNGPLEEASIALYYSTGCTQLPTLARSFTICIMISEAR 170
DB 156 GNYDRLEQLAGNLRENIELGNGPLEEASIALYYSTGCTQLPTLARSFTICIMISEAR 215
QY 171 FOYIEGEMRTIRIRNRRS 188
DB 216 FOYIEGEMRTIRIRNRRS 233

RESULT 7
AAP95639
ID AAP95639 standard; protein; 332 AA.
XX AC AAP95639;
XX DT 25-MAR-2003 (updated)
XX DT 31-OCT-2002 (updated)

```



QY	171	FQYEGEMRTIRYNRRS	188
Db	216	FQYEGEMRTIRYNRRS	233
RESULT 8			
ID	AA070827	standard; Protein; 554 AA.	
XX	AA070827;		
AC	25-MAR-2003	(updated)	
DT	31-AUG-1995	(first entry)	
XX			
DE	Anti-cataract immunotoxin.		
XX			
KW	Immunotoxin; heavy chain; light chain; variable region; antibody;		
KW	ricin-A; cytosol; cataract; lens opacification; epithelial cell;		
KW	PHB19; 4197X; monoclonal antibody; Mab.		
XX			
OS	Synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	Peptide	1..27	
FT		/label= Sig_peptide	
FT		/note= "pho signal sequence"	
FT	Domain	28..145	
FT		/label= HEAVY	
FT		/note= "Mab 4197X heavy chain"	
FT	Peptide	148..166	
FT		/label= LINER	
FT	Domain	169..274	
FT		/label= LIGHT	
FT		/note= "Mab 419X light chain"	
FT	Domain	276..544	
FT		/label= RICIN-A	
FT	Peptide	549..554	
FT		/label= TAG	
FT		/note= "hexa-histidine tail"	
XX			
XX	W09503828-A1.		
XX			
PD	09-FEB-1995.		
PE	15-JUL-1994;	94WO-US07919.	
XX			
PR	02-AUG-1993;	93OS-0101329.	
XX			
PA	(HOUS-) HOUSTON BIOTECHNOLOGY INC.		
XX			
PI	Gould RM, Kelleher PJ, Wallace TL, Wood MS;		
XX			
PR	WPI; 1995-082036/11.		
XX	N-PSDB; AA085386.		
XX			
PT	New single chain immunotoxin - binds specifically to epithelial		
PT	cells, for inhibiting development of sec. cataracts after		
PT	ectria; capsular cataract extraction.		
XX			
PS	Disclosure; Fig.4; 68pp; English.		
XX			
CC	The immunotoxin given in AA070827 comprises the heavy and light chain		
CC	variable regions of anti-lens epithelium 16G3 Mab 4197X linked to		
CC	rich-A and a hexa-histidine tag. The DNA construct encoding the		
CC	immunotoxin was expressed from PHB19 in E. coli.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
XX			
SO	Sequence	554 AA;	
Query Match		98.4%; Score 951; DB 16; Length 554;	
Best Local Similarity		94.9%; Pred. No. 2, 5e-93;	
Matches 188; Conservative		0; Mismatches 0; Indels 10; Gaps 1;	

```

OY 1 IFPKOYPIINTAGATVOSTNFIKRAVGRLT-----VLPNRVGLPINORFTLV 50
DB 278 IFPKOYPIINTAGATVOSTNFIKRAVGRLTGADYRHRIPVLPNRVGLPINORFTLV 337
OY 51 ELSNHAELSVTLADVTNAVYVGYRAGNSAYFFHPDNOEDAEALITHLFTDVONRYTFAFG 110
DB 338 ELSNHAELSVTLADVTNAVYVGYRAGNSAYFFHPDNOEDAEALITHLFTDVONRYTFAFG 397
OY 111 GNYDRLEQLAGNLRNIEELGNGPLEEALISALYYSTGTQLPPLARSFTICIQMISEAR 170
DB 398 GNYDRLEQLAGNLRNIEELGNGPLEEALISALYYSTGTQLPPLARSFTICIQMISEAR 457
OY 171 FOYIEGEMRTIRIRYNRRS 188
DB 458 FOYIEGEMRTIRIRYNRRS 475

RESULT 9
AAP90079 standard; protein; 562 AA.
ID AAP90079
XX AAP90079;
AC 25-MAR-2003 (updated)
DT 01-NOV-1989 (first entry)
XX 25-MAR-2003 (first entry)
DE Ricin D.
XX Ricin D; Ricinus communis; castor beans; Zanthariensis variety;
KW modified; lectin binding removed; reduced cell binding
XX Ricinus communis (caster beans).
OS
XX W08904839-A.
PN 01-JUN-1989.
XX 23-NOV-1988; 88MO-US04238.
XX 24-NOV-1987; 87US-0124735.
PR (GEMY) GENETICS INST INC.
XX Brown EL, Jones S;
PI WPI, 1985-178366/24.
DR N-PSDB; AAN90068.
XX Modified ricin molecules and toxin conjugates
PT - in which the lectin binding function of the B chain
PT is removed or diminished to reduce cell binding.
XX Disclosure; fig 1; 51pp; English.
XX Ricin D (see corresp. AAN90068) found on the EcoRI-HindIII fragment
CC of DNA from Ricinu communis, zanthariensis variety. Patent
CC discloses many modifications of ricin in which the lectin binding
CC function of the B chain is diminished or removed, and conjugation
CC to toxins to eliminate cell binding
CC (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX Sequence 562 AA;
SO
Query Match 98.4%; Score 951; DB 10; Length 562;
Best Local Similarity 94.9%; Pred. No. 2.5e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
OY 1 IFPKOYPIINTAGATVOSTNFIKRAVGRLT-----VLPNRVGLPINORFTLV 50
DB 36 IFPKOYPIINTAGATVOSTNFIKRAVGRLTGADYRHRIPVLPNRVGLPINORFTLV 95
OY 51 ELSNHAELSVTLADVTNAVYVGYRAGNSAYFFHPDNOEDAEALITHLFTDVONRYTFAFG 110

```

```

DB 96 ELSNHAELSVTLADVTNAVYVGYRAGNSAYFFHPDNOEDAEALITHLFTDVONRYTFAFG 155
OY 111 GNYDRLEQLAGNLRNIEELGNGPLEEALISALYYSTGTQLPPLARSFTICIQMISEAR 170
DB 156 GNYDRLEQLAGNLRNIEELGNGPLEEALISALYYSTGTQLPPLARSFTICIQMISEAR 215
OY 171 FOYIEGEMRTIRIRYNRRS 188
DB 216 FOYIEGEMRTIRIRYNRRS 233

RESULT 10
AAP50166 standard; Protein; 565 AA.
ID AAP50166
XX AAP50166;
AC 16-OCT-1991 (first entry)
DT Sequence of preporicin encoded by pRCL617.
XX Sequence of preporicin encoded by pRCL617.
DE Toxin; anti-tumour therapy.
XX Ricinus.
OS
XX Key Location/Qualifiers
FH 1..24
FH Peptide /label= signal
FT 25..565
FT Protein /label= N-linked glycosylation
FT 292..303
FT Region /label= links the C-terminus of the A chain and
FT the N-terminus of the B chain
FT Modified-site 34..36
FT /label= N-linked glycosylation
FT Modified-site 260..262
FT /label= N-linked glycosylation
FT Modified-site 398..400
FT /label= N-linked glycosylation
FT Modified-site 438..440
FT /label= N-linked glycosylation
PN EP145111-A.
XX 19-JUN-1985.
PD 13-JUL-1984; 84EP-0304801.
XX 13-MAR-1984; 84GB-0006569.
PR 15-JUL-1983; 83GB-0019265.
PR 15-JUL-1983; 83CH-0019265.
XX (UYWA-) UNIV WARWICK.
XX Lord JM, Roberts LM, Lamb FI;
PI WPI, 1985-148040/25.
XX N-PSDB; AAN50202.
DR New DNA sequences coding for ricin type plant toxin - or its
XX mutants, and modified vectors and host microorganisms
XX Disclosure; Page 30-30c; 40pp; English.
XX Preporicin is the whole polypeptide encoded by AAN50202 and the DNA
CC encoding this is claimed. Preporicin is obtained from preporicin by
CC removal of the AA leader sequence. The linker AA sequence which is
CC present in the precursor polypeptide is enzymatically removed in the
CC cell to separate the A and B chains, which are joined by a
CC disulphide bridge during the formation of the ricin molecule itself.
CC This linker region as well as the presumptive amino terminal leader
CC or signal sequence are not present in the sequences already
CC published by Funatsu et al.

```

XX Sequence 565 AA;  
 SQ Query Match 98.4%; Score 951; DB 6; Length 565;  
 Best Local Similarity 94.9%; Pred. No. 2.5e-93;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFFKQYPIINFTAGATVOSTYNTFIRAVRGRLT-----VLPRVGLPINOREFLV 50  
 DB 25 IFFKQYPIINFTAGATVOSTYNTFIRAVRGRLTGGADVRHDIPLVPRVGLPINOREFLV 84  
 QY 51 ELSNHAELSVTLALDVTNNAVYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVQNRYPFAFG 110  
 DB 85 ELSNHAELSVTLALDVTNNAVYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVQNRYPFAFG 144  
 QY 111 GNYDLRLEQLAGNLRENIETLGNGLPELAEISALYYSTGCTQLPTLARSFIICOMISEAR 170  
 DB 145 GNYDLRLEQLAGNLRENIETLGNGLPELAEISALYYSTGCTQLPTLARSFIICOMISEAR 204  
 QY 171 FOYIEGEMRTIRIRYNNRS 188  
 DB 205 FOYIEGEMRTIRIRYNNRS 222

RESULT 11  
 AAG78300 ID AAG78300 standard; Protein; 565 AA.  
 XX AAG78300;  
 XX  
 DT 15-NOV-2001 (first entry)  
 XX  
 DE Castor bean preproricin protein (SEQ ID 1).  
 KM Castor bean plant; preproricin; ricin; A chain; B chain;  
 KM human immunodeficiency virus infection; HIV; toxin; antiviral agent;  
 KM retroviral infection; anti-HIV; virucide activity; viral protease.  
 XX Ricinus communis.  
 OS  
 FH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT Protein /label= Signal peptide  
 FT 25..290  
 FT /label= Ricin\_A\_chain  
 FT /note= "N-glycosidase"  
 FT 291..302  
 FT /label= Linker\_peptide  
 FT /note= "Cleaved during activation of ricin"  
 FT Protein 303..565  
 FT /label= Ricin\_B\_chain  
 FT /note= "Galactose/N-acetylgalactosamine-binding lectin"  
 FT  
 PN WO200160393-A1.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 15-FEB-2001; 2001WO-US05282.  
 XX  
 PR 16-FEB-2000; 2000US-0182759.  
 XX  
 PA (BECH-) BECHTEL BKXT IDAHO LLC.  
 XX  
 PI Keener WK, Ward TE;  
 XX  
 DR WPI; 2001-581908/65.  
 DR N-PSDB; AAI64137.  
 XX  
 PT Novel composition comprising toxin e.g., ricin based antiviral compound  
 PT useful for treating viral infections such as human immunodeficiency  
 PT virus infection.  
 XX  
 PS Disclosure: Page 47-50; 66pp; English.

XX The sequence relates to preproricin protein encoded by the DNA sequence  
 CC given in AAI64137. The invention relates to a novel toxin (e.g., ricin)  
 CC based antiviral agent which is toxic to virus-infected cells, but  
 CC non-toxic to uninfected cells. The invention has anti-HIV and virucide  
 CC activities. Its mechanism of action is through inactivation of cellular  
 CC ribosomes and enhancement of binding of the antiviral agent to galactose  
 CC residues on cell surfaces, and its cellular internalisation. The  
 CC invention is useful for treating human immunodeficiency virus infection  
 CC and other viral infections, especially retroviral infections. The  
 CC antiviral agent is activated in viral particles or early-stage infected  
 CC cells, killing the cells upon infection and effectively preventing the  
 CC integration of the viral genome into the host genome thereby preventing  
 CC the latency/rebound problem. The agent enters all HIV susceptible cells,  
 CC and not just cells known to act as host cells for the virus. The  
 CC antiviral agent remains inert in a cell until degraded in it, unless the  
 CC cell is infected with the virus, where the viral protease activates it.

SQ Sequence 565 AA;  
 QY Query Match 98.4%; Score 951; DB 22; Length 565;  
 DB Best Local Similarity 94.9%; Pred. No. 2.5e-93;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFFKQYPIINFTAGATVOSTYNTFIRAVRGRLT-----VLPRVGLPINOREFLV 50  
 DB 25 IFFKQYPIINFTAGATVOSTYNTFIRAVRGRLTGGADVRHDIPLVPRVGLPINOREFLV 84  
 QY 51 ELSNHAELSVTLALDVTNNAVYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVQNRYPFAFG 110  
 DB 85 ELSNHAELSVTLALDVTNNAVYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVQNRYPFAFG 144  
 QY 111 GNYDLRLEQLAGNLRENIETLGNGLPELAEISALYYSTGCTQLPTLARSFIICOMISEAR 170  
 DB 145 GNYDLRLEQLAGNLRENIETLGNGLPELAEISALYYSTGCTQLPTLARSFIICOMISEAR 204  
 QY 171 FOYIEGEMRTIRIRYNNRS 188  
 DB 205 FOYIEGEMRTIRIRYNNRS 222

RESULT 12  
 AAG78304 ID AAG78304 standard; Protein; 565 AA.  
 XX AAG78304;  
 XX  
 AC AAG78304;  
 XX  
 DT 27-NOV-2001 (first entry)  
 XX  
 DE Modified castor bean preproricin (SEQ ID 10).  
 XX  
 DE Castor bean plant; preproricin; ricin; A chain; B chain;  
 KM human immunodeficiency virus infection; HIV; toxin; antiviral agent;  
 KM retroviral infection; anti-HIV; virucide; viral protease.  
 XX  
 OS Chimeric - Ricinus communis  
 OS Chimeric - Human immunodeficiency virus type 2.  
 FH  
 FH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT Protein /label= Signal\_peptide  
 FT 25..565  
 FT /label= Proricin  
 FT /note= "Proricin consists of the ricin A chain, a linker  
 FT peptide, and the ricin B chain. Proricin is  
 FT proteolytically cleaved between the A chain and  
 FT the linker to yield mature ricin"  
 FT Protein 25..291  
 FT /label= Ricin\_A\_chain  
 FT /note= "N-glycosidase"  
 FT 292..303  
 FT /label= Linker\_peptide  
 FT 296..297  
 FT Cleavage-site



DB 156 GNYDRLEQLAGNLRNIELGNGPLEEALISALYYSTGTGOLPTLARSFICIQMISEAR 215  
 QY 171 FOYIEGEMRTIRYRNR 188  
 |||||  
 DB 216 FOYIEGEMRTIRYRNR 233

RESULT 14  
 AAW25787  
 ID AAW25787 standard; Protein: 576 AA.  
 XX  
 AC AAW25787;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 27-MAR-1998 (first entry)  
 XX  
 DE Castorbean ricin.  
 XX  
 KW Ricin; cytotoxin; hybrid protein; cell delivery;  
 KW cell binding ligand; translocation domain; diphtheria toxin B';  
 KW Interleukin-2; T-cell lymphoma; organ rejection; therapy.  
 XX  
 OS Ricinus communis.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..35  
 FT Protein /label= sig\_peptide  
 FT /label= A-domain  
 FT Peptide 303..314  
 FT /label= Linker  
 FT Domain 315..576  
 FT /label= B-domain  
 XX  
 XX US5668255-A.  
 XX  
 PD 16-SEP-1997.  
 XX  
 PE 04-AUG-1993; 93US-0102387.  
 XX  
 PR 27-JUN-1991; 91US-0722484.  
 PR 07-JUN-1984; 84US-0618199.  
 PR 25-APR-1985; 85US-0726808.  
 PR 07-JUN-1985; 85US-0742554.  
 PR 22-DEC-1989; 89US-0456095.  
 PR 14-JUN-1990; 90US-0538276.  
 PR 04-AUG-1993; 93US-0102387.  
 XX  
 PA (SERA-) SERAGEN INC.  
 XX  
 PI Murphy JR;  
 XX  
 DR WPI; 1997-470103/43.  
 DR N-PSDB; AAT91638.  
 XX  
 XX New hybrid molecules for delivery of agents to cells - comprise a  
 PT binding domain of a cell binding ligand and a portion of a  
 PT translocation domain of a protein  
 XX  
 XX Example 4; Fig 11A-B; 30pp; English.  
 XX  
 XX This polypeptide comprises the castorbean cytotoxin, ricin.  
 CC DNA (see AAT91638) encoding the enzymatic A domain and a portion  
 CC of the A-to-B linker peptide of ricin was used to construct a  
 CC ricin-diphtheria toxin B-interleukin-2 gene that was expressed in  
 CC E. coli. The hybrid protein can be isolated and used to treat  
 CC conditions involving over-production of cells bearing IL2 receptors,  
 CC such as certain T-cell lymphomas and organ transplant rejection  
 CC crises. The hybrid inactivates ribosomes in cells bearing IL2  
 CC receptors, resulting in cessation of protein synthesis and death of  
 CC target cells. Claimed hybrid proteins comprise a translocation  
 CC domain and a cell binding domain from e.g. a hormone, growth factor  
 CC or polypeptide toxin. The hybrid molecules can be used for the

CC delivery of agents (e.g. therapeutic genes, toxins, detectable  
 CC labels) into cells. The use of a translocation mechanism ensures  
 CC that the hybrid will be effective in relatively low doses, since a  
 CC high proportion of the substance of interest will be taken into the  
 CC targeted cells. The hybrid molecules can be manufactured as a  
 CC single hybrid recombinant protein, permitting reproducibility,  
 CC consistency, and the precise control of composition.  
 CC (updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 576 AA;  
 XX  
 Query Match 98.4%; Score 951; DB 18; Length 576;  
 Best Local Similarity 94.9%; Pred. No. 2.6e-93;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IEPKOYPIINFTAGATVQSYTNFTRAVNGRLT-----VLPNRVGLPIINQFTILV 50  
 |||||  
 DB 36 IEPKOYPIINFTAGATVQSYTNFTRAVNGRLTGVADYRHEILVLPNRVGLPIINQFTILV 95  
 |||||  
 QY 51 ELSNHAELSVTLADVTNAYVYVGRAGNSAYFFHNDQEDAEATHTLFTDVONRYTFAFG 110  
 |||||  
 DB 96 ELSNHAELSVTLADVTNAYVYVGRAGNSAYFFHNDQEDAEATHTLFTDVONRYTFAFG 155  
 |||||

QY 111 GNYDRLEQLAGNLRNIELGNGPLEEALISALYYSTGTGOLPTLARSFICIQMISEAR 170  
 |||||  
 DB 156 GNYDRLEQLAGNLRNIELGNGPLEEALISALYYSTGTGOLPTLARSFICIQMISEAR 215  
 |||||

QY 171 FOYIEGEMRTIRYRNR 188  
 |||||  
 DB 216 FOYIEGEMRTIRYRNR 233  
 |||||

RESULT 15  
 AAY55892  
 ID AAY55892 standard; Protein: 576 AA.  
 XX  
 AC AAY55892;  
 XX  
 DT 15-FEB-2000 (first entry)  
 DT  
 XX  
 DE Castor bean ricin toxin.  
 XX  
 KW Recombinant; hybrid; binding domain; ligand; animal cell; diphtheria;  
 KW translocation domain; botulinum; neurotoxin; ricin; cholera; tetanus;  
 KW shiga-like toxin; pertussis; translocation; cytoplasmic membrane; HIV;  
 KW cytosol; therapy; genetic deficiency disease; enzyme; co-factor; poison;  
 KW adipocyte; cancer; virus; infection; antibody.  
 XX  
 XX Ricinus communis.  
 OS  
 XX  
 PN US5965406-A.  
 XX  
 PD 12-OCT-1999.  
 XX  
 PE 07-JUN-1995; 95US-0488246.  
 XX  
 PR 04-AUG-1993; 93US-0102387.  
 PR 07-JUN-1984; 84US-0618199.  
 PR 27-JUN-1991; 91US-0722484.  
 PR 25-APR-1985; 85US-0726808.  
 PR 07-JUN-1985; 85US-0742554.  
 PR 22-DEC-1989; 89US-0456095.  
 PR 14-JUN-1990; 90US-0538276.  
 XX  
 PA (SERA-) SERAGEN INC.  
 XX  
 PI Murphy JR;  
 XX  
 DR WPI; 1999-632431/54.  
 DR N-PSDB; AAZ30663.  
 XX  
 PT Recombinant DNA molecule encoding a three part hybrid protein used in  
 the treatment of Aids and genetic deficiency diseases

XX Example 4; Fig 11: 31pp: English.

PS  
XX  
CC The invention relates to a recombinant DNA molecule encoding a hybrid  
CC protein comprising three parts: (a) the first part comprises a portion  
CC of the binding domain of a cell-binding polypeptide ligand allowing the  
CC hybrid protein to bind to an animal cell; (b) the second part comprises  
CC a portion of a translocation domain of a naturally occurring protein  
CC selected from diphtheria toxin, Botulinum neurotoxin, ricin, cholera  
CC toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus  
CC toxin, which translocate the third part of the across the cytoplasmic  
CC membrane into the cytosol of the cell; and (c) the third part comprises  
CC a polypeptide entity to be introduced into the cell, which is non-native  
CC to the naturally occurring protein of (b). This sequence represents the  
CC Castor bean ricin toxin sequence for use in generating the hybrid of the  
CC invention. The hybrid molecule enables the direction of appropriate  
CC therapy to affected cells, allowing them to function properly and  
CC alleviate or cure the disease. The hybrid is especially used in treating  
CC genetic deficiency diseases, by delivering to affected cells an enzyme  
CC supplying the missing function, to supplementing cellular levels of a  
CC particular enzyme or a scarce precursor or cofactor, to directing toxins  
CC or other poisons to destroy particular cells (such as adipocytes, cancer  
CC cell, or virus infected-cells), to counteracting viral infections such as  
CC HIV, by introducing appropriate antibodies to viral proteins. It is also  
CC involved in the process of getting non-therapeutic substances such as  
CC detectable labels into cells.

XX  
SQ Sequence 576 AA:

Query Match 98.4%; Score 951; DB 20; Length 576;

Best Local Similarity 94.9%; Pred. No. 2.6e-93;

Matches 186; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFFKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILY 50  
DB 36 IFFKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILY 95  
QY 51 ELSNHAELSTLADVTNNAVVGYSRAGNSAYFFHPNODAEATHLFTDVONRYTFAG 110  
DB 96 ELSNHAELSTLADVTNNAVVGYSRAGNSAYFFHPNODAEATHLFTDVONRYTFAG 155  
QY 111 GNYDRLEQLAGNLRENIEELGNGPLFEAISALYYSTGQTOLPTLARSFTICIMISEAR 170  
DB 156 GNYDRLEQLAGNLRENIEELGNGPLFEAISALYYSTGQTOLPTLARSFTICIMISEAR 215  
QY 171 FQYIEGEMRTIRYRRS 188  
DB 216 FQYIEGEMRTIRYRRS 233

Search completed: September 16, 2003, 11:45:17  
Job time : 31.3059 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:42:03 ; Search time 10.7099 Seconds  
(without alignments)  
742.718 Million cell updates/sec

Title: US-10-083-336a-4

Perfect score: 966  
Sequence: 1 IEPKQYPIINFTTAGATVOS.....ARFOYIEGEMRTIRYRNRSS 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	951	98.4	267	1	US-07-901-707-1
2	951	98.4	267	1	US-07-988-430-1
3	951	98.4	267	1	US-08-425-336-1
4	951	98.4	267	1	US-08-488-113B-1
5	951	98.4	267	1	US-08-477-484B-1
6	951	98.4	267	1	US-08-646-360-1
7	951	98.4	267	3	US-08-839-765-1
8	951	98.4	267	3	US-09-136-389-1
9	951	98.4	267	5	US-09-610-838-1
10	951	98.4	267	5	PCT-US92-09487-1
11	951	98.4	267	2	US-08-356-786-8
12	951	98.4	267	2	US-08-378-761A-27
13	951	98.4	267	2	US-08-485-286-27
14	951	98.4	267	6	5248606-4
15	951	98.4	267	2	US-08-356-786-10
16	951	98.4	267	1	US-08-218-303-15
17	951	98.4	267	2	US-08-338-792D-61
18	951	98.4	267	4	US-09-538-873-1
19	951	98.4	267	4	US-08-378-761A-77
20	861.5	89.2	540	1	US-08-485-286-77
21	861.5	89.2	540	1	US-08-488-113B-6
22	336	34.8	247	2	US-08-477-484B-6
23	336	34.8	247	2	US-08-646-360-6
24	336	34.8	247	3	US-08-839-765-6
25	336	34.8	247	3	US-09-136-389-6
26	336	34.8	247	4	US-09-610-838-6
27	336	34.8	267	1	US-08-378-761A-74

28	336	34.8	267	1	US-08-485-286-74	Sequence 74, Appl
29	336	34.8	289	1	US-07-923-692C-4	Sequence 4, Appl
30	336	34.8	289	1	US-08-184-237-4	Sequence 4, Appl
31	336	34.8	289	2	US-08-482-920-4	Sequence 4, Appl
32	336	34.8	289	3	US-08-484-341-4	Sequence 4, Appl
33	336	34.8	289	3	US-08-483-502-4	Sequence 4, Appl
34	336	34.8	289	4	US-09-726-651A-4	Sequence 4, Appl
35	320.5	33.2	282	1	US-08-324-301-15	Sequence 15, Appl
36	310.5	32.1	255	1	US-07-901-707-6	Sequence 6, Appl
37	310.5	32.1	255	1	US-07-988-430-6	Sequence 6, Appl
38	310.5	32.1	255	1	US-08-425-336-6	Sequence 6, Appl
39	310.5	32.1	255	5	PCT-US92-09487-6	Sequence 6, Appl
40	305	31.6	248	3	US-08-902-486-7	Sequence 7, Appl
41	305	31.6	290	1	US-08-245-754A-2	Sequence 2, Appl
42	305	31.6	290	2	US-08-597-731-2	Sequence 2, Appl
43	305	31.6	496	3	US-08-902-486-15	Sequence 15, Appl
44	303.5	31.4	250	1	US-08-378-761A-71	Sequence 71, Appl
45	303.5	31.4	250	1	US-08-485-286-71	Sequence 71, Appl

#### ALIGNMENTS

RESULT 1  
US-07-901-707-1  
Sequence 1, Application US/07901707  
Patent No. 5376546  
GENERAL INFORMATION:  
APPLICANT: Bernhardt, Susan L.  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Steve F.  
APPLICANT: Lane, Julie A.  
TITLE OF INVENTION: Materials Comprising and Methods of  
TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/901,707  
FILING DATE: 19920619  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: 5376546and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27129/30910  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-5750  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-901-707-1  
Query Match 98.4% Score 951, DB 1, Length 267;

Best Local Similarity 94.9%; Pred. No. 1.8e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
Db

1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLNRYGLPINOFTILV 50  
1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGVADVRHEIPLVLPNRGGLPINORFTILV 60  
51 ELSNHAELSVTLALDVYNAVYVGRAGNSAYFFHPDNOEDAEATITLFTDVONRYTFAPG 110  
61 ELSNHAELSVTLALDVYNAVYVGRAGNSAYFFHPDNOEDAEATITLFTDVONRYTFAPG 120  
111 GNDRIEQLAGNLRENIEELNGPLLEAISAALYYSTGTQPLPTLARSFTICMISEAR 170  
121 GNDRIEQLAGNLRENIEELNGPLLEAISAALYYSTGTQPLPTLARSFTICMISEAR 180  
171 FQYIEGEMRTIRIRYNNRS 188  
181 FQYIEGEMRTIRIRYNNRS 198

RESULT 2  
US-07-988-430-1  
Sequence 1, Application US/07988430  
Patent No. 5416202

## GENERAL INFORMATION:

APPLICANT: Bernhard, Susan L.  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Lane, Julie A.  
APPLICANT: Lei, Shau-Ping  
TITLE OF INVENTION: Materials Comprising and Methods of  
Preparation and Use for Ribosome-Inactivating Proteins  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/988,430  
FILING DATE: 19921209  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:  
NAME: No. 5416202and, Greta E.

REGISTRATION NUMBER: 35302  
REFERENCE/DOCKET NUMBER: 31133

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740

TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID

TOPOLOGY: linear  
MOLECULE TYPE: protein

US-07-988-430-1

Query Match 98.4%; Score 951; DB 1; Length 267;  
Best Local Similarity 94.9%; Pred. No. 1.8e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
Db

1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLNRYGLPINOFTILV 50  
1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGVADVRHEIPLVLPNRGGLPINORFTILV 60  
51 ELSNHAELSVTLALDVYNAVYVGRAGNSAYFFHPDNOEDAEATITLFTDVONRYTFAPG 110  
61 ELSNHAELSVTLALDVYNAVYVGRAGNSAYFFHPDNOEDAEATITLFTDVONRYTFAPG 120  
111 GNDRIEQLAGNLRENIEELNGPLLEAISAALYYSTGTQPLPTLARSFTICMISEAR 170  
121 GNDRIEQLAGNLRENIEELNGPLLEAISAALYYSTGTQPLPTLARSFTICMISEAR 180  
171 FQYIEGEMRTIRIRYNNRS 188  
181 FQYIEGEMRTIRIRYNNRS 198

RESULT 3  
US-08-425-336-1  
Sequence 1, Application US/08425336  
Patent No. 5621083

## GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studlika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/425,336  
FILING DATE: 18-APR-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064,691  
FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Thomas C.

REGISTRATION NUMBER: P-36,989  
REFERENCE/DOCKET NUMBER: 31394

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448

TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-425-336-1

Query Match 98.4%; Score 951; DB 1; Length 267;



Best Local Similarity 94.9%; Pred. No. 1.8e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 1 IFPKOYPIINFTAGATVOSTYNTFIRAVRGRLT-----VLNPNVGLPINORFTLV 50  
Db 1 IFPKOYPIINFTAGATVOSTYNTFIRAVRGRLTGGADVREHIEPVLPNVRGLPINORFTLV 60  
QY 51 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATITLFTDVONRTPFAG 110  
Db 61 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATITLFTDVONRTPFAG 120  
QY 111 GNYDRLEQLAGNLRENIEELGNGLPEEALISALYYSTGCTQLPTLARSFIIQIMISEAR 170  
Db 121 GNYDRLEQLAGNLRENIEELGNGLPEEALISALYYSTGCTQLPTLARSFIIQIMISEAR 180  
QY 171 FOYIEGEMRTIRIRYNRRS 188  
Db 181 FOYIEGEMRTIRIRYNRRS 198  
RESULT 4  
US-08-488-113B-1  
; Sequence 1, Application US/08488113B  
; Patent No. 5744580  
; GENERAL INFORMATION:  
; APPLICANT: Bectel, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studulka, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Mcandrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,113B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-488-113B-1  
Query Match  
Best Local Similarity 98.4%; Score 951; DB 1; Length 267;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 1 IFPKOYPIINFTAGATVOSTYNTFIRAVRGRLT-----VLNPNVGLPINORFTLV 50  
Db 1 IFPKOYPIINFTAGATVOSTYNTFIRAVRGRLTGGADVREHIEPVLPNVRGLPINORFTLV 60  
QY 51 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATITLFTDVONRTPFAG 110  
Db 61 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATITLFTDVONRTPFAG 120  
QY 111 GNYDRLEQLAGNLRENIEELGNGLPEEALISALYYSTGCTQLPTLARSFIIQIMISEAR 170  
Db 121 GNYDRLEQLAGNLRENIEELGNGLPEEALISALYYSTGCTQLPTLARSFIIQIMISEAR 180  
QY 171 FOYIEGEMRTIRIRYNRRS 188  
Db 181 FOYIEGEMRTIRIRYNRRS 198  
RESULT 5  
US-08-477-484B-1  
; Sequence 1, Application US/08477484B  
; Patent No. 5756699  
; GENERAL INFORMATION:  
; APPLICANT: Bectel, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studulka, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Mcandrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,484B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-477-484B-1

Query Match 98.4%; Score 951; DB 1; Length 267;  
Best Local Similarity 94.9%; Pred. No. 1.8e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORITLY 50  
DB 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTGTGADVREHIFVLPNRVGLPINORITLY 60  
QY 51 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAG 110  
DB 61 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAG 120  
QY 111 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYSTGCTQLPTLARSFIIQIMISEAR 170  
DB 121 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYSTGCTQLPTLARSFIIQIMISEAR 180  
QY 171 FOYIEGEMTRIRYNRRS 188  
DB 181 FOYIEGEMTRIRYNRRS 198

## RESULT 6

US-08-646-360-1  
Sequence 1, Application US/08646360  
Patent No. 5837491  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.14  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-360-1

Query Match 98.4%; Score 951; DB 2; Length 267;  
Best Local Similarity 94.9%; Pred. No. 1.8e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORITLY 50  
DB 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTGTGADVREHIFVLPNRVGLPINORITLY 60  
QY 51 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAG 110  
DB 61 ELSNHAELSVTLALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVONRYTFAG 120  
QY 111 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYSTGCTQLPTLARSFIIQIMISEAR 170  
DB 121 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYSTGCTQLPTLARSFIIQIMISEAR 180  
QY 171 FOYIEGEMTRIRYNRRS 188  
DB 181 FOYIEGEMTRIRYNRRS 198

## RESULT 7

US-08-839-765-1  
Sequence 1, Application US/08839765  
Patent No. 6146631  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,765  
FILING DATE: 15-APR-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-839-765-1

Query Match 98.4%; Score 951; DB 3; Length 267;  
Best Local Similarity 94.9%; Pred. No. 1.8e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFFPKQPIINFTTAGATVOSTNTFRVGRGLT-----VLPNRVGLPINORFIV 50  
DB 1 IFFPKQPIINFTTAGATVOSTNTFRVGRGLT-----VLPNRVGLPINORFIV 60  
QY 51 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAFG 110  
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAFG 120  
QY 111 GNYDRLEQLAGNLRENIELGNGPLEAISALYYSTGCTQLPTLARSFTICIMISEAR 170  
DB 121 GNYDRLEQLAGNLRENIELGNGPLEAISALYYSTGCTQLPTLARSFTICIMISEAR 180  
QY 171 FOYIEGEMTRIRIRYNRS 188  
DB 181 FOYIEGEMTRIRIRYNRS 198

RESULT 8  
US-09-136-389-1  
Sequence 1, Application US/09136389  
Patent No. 6146850  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,389  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/646,360

FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-136-389-1

Query Match 98.4%; Score 951; DB 3; Length 267;  
Best Local Similarity 94.9%; Pred. No. 1.8e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFFPKQPIINFTTAGATVOSTNTFRVGRGLT-----VLPNRVGLPINORFIV 50  
DB 1 IFFPKQPIINFTTAGATVOSTNTFRVGRGLT-----VLPNRVGLPINORFIV 60  
QY 51 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAFG 110  
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVONRYTFAFG 120  
QY 111 GNYDRLEQLAGNLRENIELGNGPLEAISALYYSTGCTQLPTLARSFTICIMISEAR 170  
DB 121 GNYDRLEQLAGNLRENIELGNGPLEAISALYYSTGCTQLPTLARSFTICIMISEAR 180  
QY 171 FOYIEGEMTRIRIRYNRS 188  
DB 181 FOYIEGEMTRIRIRYNRS 198

RESULT 9  
US-09-610-838-1  
Sequence 1, Application US/09610838  
Patent No. 6376217  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible



NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356,786  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/831,967  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-053  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-356-786-8

Query Match  
Best Local Similarity 98.4%; Score 951; DB 2; Length 268;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLT-----VLPRVGLPIINQRIIV 50  
DB 2 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPIINQRIIV 61  
QY 51 ELSNHAELSVTLALDVTNAVYVGYRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 110  
DB 62 ELSNHAELSVTLALDVTNAVYVGYRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 121  
QY 111 GNYDRLEQLAGNREINTELGNGPLEEASALYYSTGGTQLPTLARSFTICIMISEAR 170  
DB 122 GNYDRLEQLAGNREINTELGNGPLEEASALYYSTGGTQLPTLARSFTICIMISEAR 181  
QY 171 FOYTEGEMRTIRIRYNRRS 188  
DB 182 FOYTEGEMRTIRIRYNRRS 199

RESULT 12  
US-08-378-761A-27  
Sequence 27, Application US/08378761A  
Patent No. 5635384  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN

COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/378,761A  
FILING DATE: 26-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BORUCKI, ANDREA T  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 38272B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 337-4846  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 290 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-378-761A-27

Query Match  
Best Local Similarity 98.4%; Score 951; DB 1; Length 290;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLT-----VLPRVGLPIINQRIIV 50  
DB 25 IFPKQYPIINFTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPIINQRIIV 84  
QY 51 ELSNHAELSVTLALDVTNAVYVGYRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 110  
DB 85 ELSNHAELSVTLALDVTNAVYVGYRAGNSAYFFHPDNOEDAETHLFTDVONRYTFAG 144  
QY 111 GNYDRLEQLAGNREINTELGNGPLEEASALYYSTGGTQLPTLARSFTICIMISEAR 170  
DB 145 GNYDRLEQLAGNREINTELGNGPLEEASALYYSTGGTQLPTLARSFTICIMISEAR 204  
QY 171 FOYTEGEMRTIRIRYNRRS 188  
DB 205 FOYTEGEMRTIRIRYNRRS 222

RESULT 13  
US-08-485-286-27  
Sequence 27, Application US/08485286  
Patent No. 5646026  
Patent No. 5646026 5646119  
GENERAL INFORMATION:  
APPLICANT: WALSH, TERENCE A  
APPLICANT: HEY, TIMOTHY D  
APPLICANT: MORGAN, ALICE ER  
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE  
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ANDREA T. BORUCKI  
STREET: 9330 ZIONSVILLE ROAD  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: US  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:



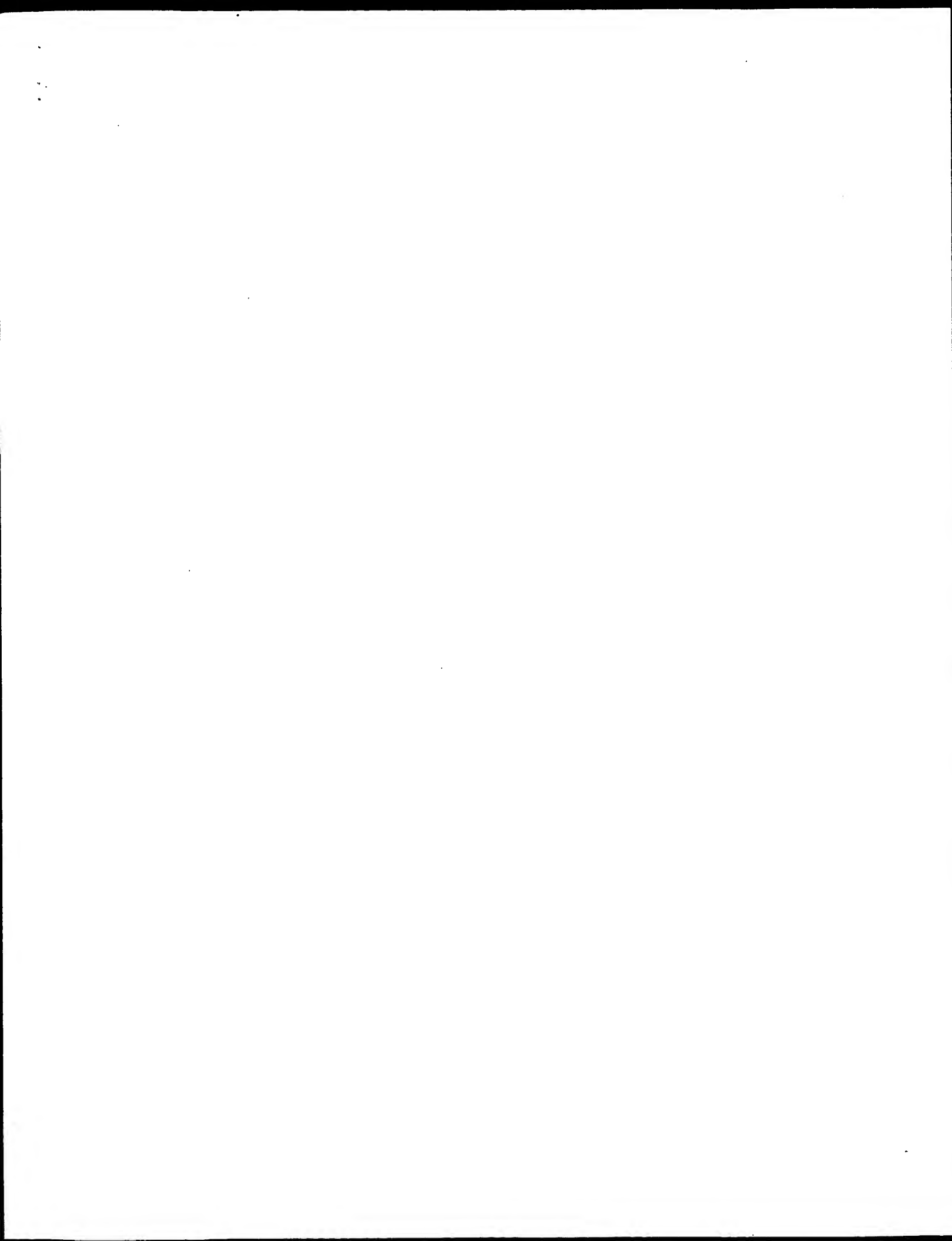
Tue Sep 16 12:27:50 2003

us-10-083-336a-4.rai

Page 9

Job time : 11.7099 secs

---





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:49:23 : Search time 19.4426 Seconds  
(without alignments)  
1439.723 Million cell updates/sec

Title: US-10-083-336a-4

Perfect score: 966

Sequence: 1 IFPKQYPIINFTAGATVQS.....ARFYIGEMRTIRIRNRS 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 556269 seqs, 148893369 residues

Total number of hits satisfying chosen parameters: 556269

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PC1\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCRTS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	951	98.4	267	12	US-10-127-890-1
2	941	97.4	267	12	US-10-282-935-1
3	336	34.8	247	10	US-09-792-793A-39
4	336	34.8	247	12	US-10-127-890-6
5	336	34.8	289	12	US-10-280-579B-4
6	305	31.6	247	10	US-09-792-793A-34
7	297.5	30.8	251	12	US-10-282-935-3
8	273	28.3	263	12	US-10-127-890-7
9	272	28.2	263	12	US-10-127-890-4
10	257.5	26.7	248	12	US-10-127-890-5
11	254	26.3	252	9	US-09-347-064-2
12	254	26.3	252	9	US-09-347-064-2
13	245.5	25.4	251	12	US-10-127-890-107
14	244.5	25.3	251	12	US-10-127-890-106
15	244.5	25.3	251	12	US-10-127-890-110

16	244.5	25.3	251	12	US-10-127-890-111	Sequence 111, App
17	243.5	25.2	251	9	US-09-765-527-247	Sequence 247, App
18	243.5	25.2	251	12	US-10-127-890-2	Sequence 2, App1
19	243.5	25.2	251	12	US-10-127-890-99	Sequence 99, App1
20	243.5	25.2	251	12	US-10-127-890-100	Sequence 100, App
21	243.5	25.2	251	12	US-10-127-890-101	Sequence 101, App
22	243.5	25.2	251	12	US-10-127-890-102	Sequence 102, App
23	243.5	25.2	251	12	US-10-127-890-103	Sequence 103, App
24	243.5	25.2	251	12	US-10-127-890-104	Sequence 104, App
25	243.5	25.2	251	12	US-10-127-890-105	Sequence 105, App
26	242.5	25.1	251	12	US-10-127-890-109	Sequence 109, App
27	242.5	25.1	293	9	US-09-765-527-259	Sequence 259, App
28	242.5	25.1	332	9	US-09-765-527-253	Sequence 253, App
29	242.5	25.1	332	9	US-09-765-527-251	Sequence 251, App
30	241.5	25.0	251	12	US-10-127-890-108	Sequence 108, App
31	172	17.8	250	10	US-09-792-793A-36	Sequence 36, App1
32	172	17.8	250	12	US-10-127-890-8	Sequence 8, App1
33	172	17.8	263	10	US-09-978-274A-4	Sequence 4, App1
34	172	17.8	314	10	US-09-978-274A-2	Sequence 2, App1
35	168.5	17.4	261	12	US-10-127-890-9	Sequence 9, App1
36	126	13.0	330	10	US-09-792-793A-82	Sequence 82, App1
37	125	12.9	332	10	US-09-792-793A-73	Sequence 73, App1
38	124	12.8	254	10	US-09-792-793A-85	Sequence 85, App1
39	124	12.8	259	12	US-10-127-890-10	Sequence 10, App1
40	124	12.8	275	10	US-09-792-793A-35	Sequence 35, App1
41	124	12.8	327	10	US-09-792-793A-79	Sequence 79, App1
42	124	12.8	332	10	US-09-792-793A-76	Sequence 76, App1
43	109	11.3	280	12	US-10-127-890-3	Sequence 3, App1
44	109	11.3	280	12	US-10-127-890-127	Sequence 127, App
45	109	11.3	280	12	US-10-127-890-128	Sequence 128, App

## ALIGNMENTS

RESULT 1  
US-10-127-890-1  
Sequence 1, Application US/10127890  
Publication No. US20030166196A1

GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnicka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating Proteins

NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-May-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-May-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-May-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-Dec-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-Jun-1992

APPLICATION NUMBER: US 07/787,567  
 FILING DATE: 04-NOV-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: McNicholas, Janet M.  
 REGISTRATION NUMBER: 32,918  
 REFERENCE/DOCKET NUMBER: 200-70.P4  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312/707-8889  
 TELEFAX: 312/707-9155  
 TELE: 650 368-1248  
 INFORMATION FOR SEQ ID NO: 1:  
 LENGTH: 267 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-10-127-890-1

Query Match 98.4%; Score 951; DB 12; Length 267;  
 Best Local Similarity 94.9%; Pred. No. 6.2e-101;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFFKQYPIINFETAGATVOSTNFIKAVGRRLT-----VLPNRVGLPINORFLIV 50  
 DB 1 IFFKQYPIINFETAGATVOSTNFIKAVGRRLTGDVVRHEIPLPNRVGLPINORFLIV 60  
 QY 51 ELSNAELSVTLALDVTNAYVGYRAGNSAYFFHPNODAEATITLFTDVONRTTFAFG 110  
 DB 61 ELSNAELSVTLALDVTNAYVGYRAGNSAYFFHPNODAEATITLFTDVONRTTFAFG 120  
 QY 111 GNDRLLEQLAGNIRENIEELGNGPLEAISALYYSTGQTQLPTLARSFTICQMISEAR 170  
 DB 121 GNDRLLEQLAGNIRENIEELGNGPLEAISALYYSTGQTQLPTLARSFTICQMISEAR 180  
 QY 171 FYIEGEMRTIRIRYNRRS 188  
 DB 181 FYIEGEMRTIRIRYNRRS 198

RESULT 2  
 US-10-282-935-1  
 Sequence 1, Application US/10282935  
 Publication No. US20030143193A1  
 GENERAL INFORMATION:  
 APPLICANT: VITETTA, ELLEN S.  
 APPLICANT: GHETTE, VICTOR F.  
 APPLICANT: SMALLSHAW, JOAN  
 APPLICANT: BALUNA, ROXANA G.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF  
 FILE REFERENCE: US2003-084US  
 CURRENT APPLICATION NUMBER: US/10/282,935  
 CURRENT FILING DATE: 2002-10-29  
 PRIOR APPLICATION NUMBER: 09/538,873  
 PRIOR FILING DATE: 2000-03-30  
 PRIOR APPLICATION NUMBER: 60/126,826  
 PRIOR FILING DATE: 1999-03-30  
 NUMBER OF SEQ ID NOS: 23  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 1  
 LENGTH: 267  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 OTHER INFORMATION: Peptide  
 US-10-282-935-1

Query Match 97.4%; Score 941; DB 12; Length 267;  
 Best Local Similarity 94.9%; Pred. No. 8.8e-100;  
 Matches 186; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 3 PKQYPIINFETAGATVOSTNFIKAVGRRLT-----VLPNRVGLPINORFLIV 52  
 DB 3 PKQYPIINFETAGATVOSTNFIKAVGRRLTGDVVRHEIPLPNRVGLPINORFLIV 62  
 QY 53 SNAELSVTLALDVTNAYVGYRAGNSAYFFHPNODAEATITLFTDVONRTTFAFGN 112  
 DB 63 SNAELSVTLALDVTNAYVGYRAGNSAYFFHPNODAEATITLFTDVONRTTFAFGN 122  
 QY 113 YDRLEQLAGNIRENIEELGNGPLEAISALYYSTGQTQLPTLARSFTICQMISEARFQ 172  
 DB 123 YDRLEQLAGNIRENIEELGNGPLEAISALYYSTGQTQLPTLARSFTICQMISEARFQ 182  
 QY 173 YIEGEMRTIRIRYNRRS 188  
 DB 183 YIEGEMRTIRIRYNRRS 198

RESULT 3  
 US-09-792-793A-39  
 Sequence 39, Application US/09792793A  
 Patent No. US20020168370A1  
 GENERAL INFORMATION:  
 APPLICANT: McDonald, John R.  
 APPLICANT: Coggin, Philip  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE  
 FILE REFERENCE: 25020-601D  
 CURRENT APPLICATION NUMBER: US/09/792,793A  
 CURRENT FILING DATE: 2001-02-22  
 NUMBER OF SEQ ID NOS: 93  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 39  
 LENGTH: 247  
 TYPE: PRT  
 ORGANISM: Trichosanthens killowii  
 US-09-792-793A-39

Query Match 34.8%; Score 336; DB 10; Length 247;  
 Best Local Similarity 38.9%; Pred. No. 2.6e-30;  
 Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY 9 INFTTAGATVOSTNFIKAVGRRLTLPN-----RVGLPINORFLIVELSNAL 58  
 DB 2 VSFRLSGATSSSYGVFISNLR--KALPNERKLYDIPLLNSLPQSORVALIHLTYADE 58  
 QY 59 SVTLALDVTNAYVGYRAGNSAYFFHPNODAEATITLFTDVONRTTFAFGNDRL 117  
 DB 59 TISVAIDVTNAYVGYRAGNSAYFFHPNODAEATITLFTDVONRTTFAFGNDRL 115  
 QY 118 QLAGNIRENIEELGNGPLEAISALYYSTGQTQLPTLARSFTICQMISEARFQ 177  
 DB 116 TAAGKIRIRNIPGLPALSATITFTFYNNAN-----SAASALMVLIQSTSEARFQ 170  
 QY 178 MTRT 182  
 DB 171 IGRV 175

RESULT 4  
 US-10-127-890-6  
 Sequence 6, Application US/10127890  
 Publication No. US20030166196A1  
 GENERAL INFORMATION:  
 APPLICANT: Better, Marc D.  
 APPLICANT: Carroll, Stephen F.  
 APPLICANT: Studnika, Gary M.  
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
 Proteins  
 NUMBER OF SEQUENCES: 173  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: McAndrews, Held & Malloy, Ltd.  
 STREET: 500 West Madison Street, 34th floor  
 CITY: Chicago

```

STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-127-890-6

Query Match
Best Local Similarity 34.8%; Score 336; DB 12; Length 247;
Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY 9 INFTTAGATVOSTYNTFIAVGRGLTVLPN-----RVGLPINORFIIIVLSNRAEL 58
DB 2 VSRRLSGATSSSYGVFISNLR---KALPNERKLYDIPILRSLPGSORIALIHILTNVAD 58
QY 59 SVTLALDVTNAYVVGVRAGNSAYFFHPDQEDA-EATHTLFTDVONRYTFAFGNVDRL 117
DB 59 TISVALDVTNAYVVGVRAGDTSYFF---NEASATEAKYVFKDMRKVTLPYSGNVERLQ 115
QY 118 QLAGNLRNIEELGNGPLEAISALYYSTGGTQLPTLARSFICIQMISEARPOYTEGE 177
DB 116 TAAGKIRENIPILG.PALDSAITTLEFYNNAN-----SAASALMVLIQSTSEARARYKFI 170
QY 178 MRRRI 182
DB 171 IGRV 175

RESULT 5
US-10-280-679B-4
Sequence 4, Application US/10280679B
Publication No. US20030150019A1
GENERAL INFORMATION:
APPLICANT: Large Scale Biology Corporation
TITLE OF INVENTION: Monopartite RNA Virus Transformation Vectors
FILE REFERENCE: LSBC-0109-US03
CURRENT APPLICATION NUMBER: US/10/280,679B
CURRENT FILING DATE: 2000-10-24

```

```

PRIOR APPLICATION NUMBER: 09/557,941
PRIOR FILING DATE: 2000-04-24
PRIOR APPLICATION NUMBER: 08/484,341
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 07/923,692
PRIOR FILING DATE: 1992-07-31
PRIOR APPLICATION NUMBER: 07/600,244
PRIOR FILING DATE: 1990-10-22
PRIOR APPLICATION NUMBER: 07/641,617
PRIOR FILING DATE: 1991-01-16
PRIOR APPLICATION NUMBER: 07/737,899
PRIOR FILING DATE: 1991-07-26
PRIOR APPLICATION NUMBER: 07/739,143
PRIOR FILING DATE: 1991-08-01
PRIOR APPLICATION NUMBER: 07/310,881
PRIOR FILING DATE: 1989-02-17
PRIOR APPLICATION NUMBER: 07/160,766
PRIOR FILING DATE: 1988-02-26
PRIOR APPLICATION NUMBER: 07/160,771
PRIOR FILING DATE: 1988-02-26
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 289
TYPE: PRT
ORGANISM: Chinese cucumber protein alpha-trichosanthin
US-10-280-679B-4

```

```

Query Match
Best Local Similarity 34.8%; Score 336; DB 12; Length 289;
Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

```

```

QY 9 INFTTAGATVOSTYNTFIAVGRGLTVLPN-----RVGLPINORFIIIVLSNRAEL 58
DB 25 VSRRLSGATSSSYGVFISNLR---KALPNERKLYDIPILRSLPGSORIALIHILTNVAD 81
QY 59 SVTLALDVTNAYVVGVRAGNSAYFFHPDQEDA-EATHTLFTDVONRYTFAFGNVDRL 117
DB 82 TISVALDVTNAYVVGVRAGDTSYFF---NEASATEAKYVFKDMRKVTLPYSGNVERLQ 138
QY 118 QLAGNLRNIEELGNGPLEAISALYYSTGGTQLPTLARSFICIQMISEARPOYTEGE 177
DB 139 TAAGKIRENIPILG.PALDSAITTLEFYNNAN-----SAASALMVLIQSTSEARARYKFI 193
QY 178 MRRRI 182
DB 194 IGRV 198

```

```

RESULT 6
US-09-792-793A-34
Sequence 34, Application US/09792793A
Patent No. US20020168370A1
GENERAL INFORMATION:
APPLICANT: McDonald, John R.
APPLICANT: Cogging, Philip
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
FILE REFERENCE: 25020-601D
CURRENT APPLICATION NUMBER: US/09/792,793A
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 93
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 34
LENGTH: 247
TYPE: PRT
ORGANISM: Bryonia dioica
US-09-792-793A-34

Query Match
Best Local Similarity 31.6%; Score 305; DB 10; Length 247;
Matches 65; Conservative 45; Mismatches 56; Indels 16; Gaps 4;

```

QY 9 INFTAGATVOSTYNTIRAVRGLTVLPNVRGLPI-----NORFIVELSNHAEISVT 61  
Db 2 VSRISLGATTTTGGVIGKIKMLREALPYERKRYVNIPLRRSSISGSGRTLLHLTNVDETIS 61  
QY 62 LADVTNAVVGVRAGNSAVFFPHDNOEDA-EATHTLFTDVONRTYFAFGNVDRLDOLA 120  
Db 62 VANDVTNVMYIMGLADVSYFF---NEASATEAKVFEDAKKVTLPYSGNYERLQTA 118  
QY 121 GNIRENIEIENGPLEAISALYYSTGGTQPLTARSFTICIMISEARFOYIEGEMRT 180  
Db 119 GKIRENIPGLPALDSAITLLYYTAS-----SAGSALLVLIOSTAESARKYFIEQOLQK 173  
QY 181 RI 182  
Db 174 RV 175

## RESULT 7

US-10-282-935-3  
Sequence 3, Application US/10282935  
Publication No. US20030143193A1  
GENERAL INFORMATION:  
APPLICANT: VITETTA, ELLEN S.  
APPLICANT: GHETTE, VICTOR F.  
APPLICANT: SMALLSHAW, JOAN  
APPLICANT: BALUNA, ROXANA G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF  
FILE REFERENCE: DTSD:884US  
CURRENT APPLICATION NUMBER: US/10/282,935  
PRIOR FILING DATE: 2002-10-29  
PRIOR APPLICATION NUMBER: 09/538,873  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: 60/126,826  
PRIOR FILING DATE: 1999-03-30  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 251  
TYPE: PRT  
ORGANISM: Atrius precatorius  
US-10-282-935-3

Query Match 30.8%; Score 297.5; DB 12; Length 251;  
Best Local Similarity 40.5%; Pred. No. 7.2e-26;  
Matches 75; Conservative 26; Mismatches 65; Indels 19; Gaps 5;

QY 9 INFTAGATVOSTYNTIRAVRGL-----TVLPNVRGLPTNORFIVELSNHAEISVT 60  
Db 5 IKFSTEGATQSQYKQFLEALRERLGLIHIDIPVLPDPTLOERNRYIVELSNSTESI 64  
QY 61 TLADVTNAVVGVRAGNSAVFFH--PDNOEDAETHLFTDVONRTYFAFGNVDRLDOLA 118  
Db 65 EVGIDVTNAVVAARACTQSIFLADAPSSASD-----YLTGT-DQHSILPFIYTGDLER 118  
QY 119 LAGNIRENIEIENGPLEAISALYYSTGGTQPLTARSFTICIMISEARFOYIEGEMRT 178  
Db 119 WAHDSRQOIPGLQALRHGIS---FFRSQGNDEKARKRTLLIVIIOWAAARRRYSINRV 175  
QY 179 RTIR 183  
Db 176 RVSI 180

## RESULT 8

US-10-127-890-7  
Sequence 7, Application US/10127890  
Publication No. US2003016196A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnicka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-9155  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-127-890-7

Query Match 28.3%; Score 273; DB 12; Length 263;  
Best Local Similarity 34.8%; Pred. No. 5e-23;  
Matches 63; Conservative 40; Mismatches 62; Indels 16; Gaps 4;

QY 9 INFTAGATVOSTYNTIRAVRGLTVLPNVRGLPI-----NORFIVELSNHAEISVT 61  
Db 2 VSRISLGADPRSYGMEIFIDRLNALPFRKRYVINIPLPSVSGAGRYLLMHLFVNDGKTT 61  
QY 62 LADVTNAVVGVRAGNSAVFFPHDNOEDA-EATHTLFTDVONRTYFAFGNVDRLDOLA 120  
Db 62 VANDVTNVMYIMGLADTSYFF---NEPAEIASQYVRDARRKITTLPYSGNYERLQTA 118  
QY 121 GNIRENIEIENGPLEAISALYYSTGGTQPLTARSFTICIMISEARFOYIEGEMRT 180  
Db 119 GKREKIPGLPALDSAITLLHYDS-----TAAAGALLVLIQTAEARKRYIEQOLQK 173  
QY 181 R 181  
Db 174 R 174

## RESULT 9

US-10-127-890-4  
Sequence 4, Application US/10127890





TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 107:  
US-10-127-890-107

Query Match 25.4%; Score 245.5; DB 12; Length 251;  
Best Local Similarity 36.5%; Pred. No. 6.8e-20;  
Matches 70; Conservative 31; Mismatches 64; Indels 27; Gaps 8;

QY 9 INFTAGATVOSTNTRFAVRGLTVLPNRVGLPINOR-----FLVELSNHAEISLV 60  
DB 5 VSFSTGATYTYVNFLELRYKLKPEGNSHGIPLLRKKCDPGKCFVLALSDNGOLA 64  
QY 61 TLALDVTNAYVGYRAGNSAYFEH--PDNQDEAFTLHFTD-VONRTFAFGNYDRLE 117  
DB 65 EIAIDVTSVYVGYVRNRSYFFKDAPD-----AAVEGLFKNTIKTR--LHFGGSYPSLE 117  
QY 118 QLAGN--LRENTELGNGPLEAISAALYYSTGTQLPTLARSFTICIQMISEARFOYTE 175  
DB 118 --GEKAYRETTDLGIEPLRIGIKIDENAIIDNYKPTETIASSLLVVIQWSEARFTFIE 174  
QY 176 GEMRT---RIR 183  
DB 175 NOIRNNFOQRIR 186

RESULT 14  
US-10-127-890-106  
Sequence 106, Application US/10127890  
Publication No. US20030166196A1  
GENERAL INFORMATION:

APPLICANT: Bectel, Marc D.  
Carroll, Stephen F.  
Studulka, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 106:  
US-10-127-890-106

Query Match 25.3%; Score 244.5; DB 12; Length 251;  
Best Local Similarity 36.5%; Pred. No. 8.9e-20;  
Matches 70; Conservative 31; Mismatches 64; Indels 27; Gaps 8;

QY 9 INFTAGATVOSTNTRFAVRGLTVLPNRVGLPINOR-----FLVELSNHAEISLV 60  
DB 5 VSFSTGATYTYVNFLELRYKLKPEGNSHGIPLLRKKCDPGKCFVLALSDNGOLA 64  
QY 61 TLALDVTNAYVGYRAGNSAYFEH--PDNQDEAFTLHFTD-VONRTFAFGNYDRLE 117  
DB 65 EIAIDVTSVYVGYVRNRSYFFKDAPD-----AAVEGLFKNTIKTR--LHFGGSYPSLE 117  
QY 118 QLAGN--LRENTELGNGPLEAISAALYYSTGTQLPTLARSFTICIQMISEARFOYTE 175  
DB 118 --GEKAYRETTDLGIEPLRIGIKIDENAIIDNYKPTETIASSLLVVIQWSEARFTFIE 174  
QY 176 GEMRT---RIR 183  
DB 175 NOIRNNFOQRIR 186

RESULT 15  
US-10-127-890-110

Sequence 110, Application US/10127890  
Publication No. US20030166196A1  
GENERAL INFORMATION:

APPLICANT: Bectel, Marc D.  
Carroll, Stephen F.  
Studulka, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992  
 APPLICATION NUMBER: US 07/901,707  
 FILING DATE: 19-JUN-1992  
 APPLICATION NUMBER: US 07/787,567  
 FILING DATE: 04-NOV-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: MCNICHOIAS, Janet M.  
 REGISTRATION NUMBER: 32,918  
 REFERENCE/DOCKET NUMBER: 200-70.P4  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312/707-8889  
 TELEFAX: 312/707-9155  
 TELEEX: 650 388-1248  
 INFORMATION FOR SEQ ID NO: 110:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 251 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SQUENCE DESCRIPTION: SEQ ID NO: 110:  
 US-10-127-890-110



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:40:32 ; Search time 12.1928 Seconds  
(without alignments)  
1482.817 Million cell updates/sec

Title: US-10-083-336a-4

Perfect score: 966

Sequence: 1 IEPKQPIINFTTAGATVOS.....ARFOYIEGEMTRIRYRNRSS 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	951	98.4	576	1	RLCSD
2	861.5	89.2	564	1	RLCSAG
3	336	34.8	289	1	RLRXT
4	327.5	33.9	528	2	S32431
5	327.5	33.9	562	2	S16022
6	326	33.7	247	2	JU0393
7	326	33.7	247	2	JU0393
8	326	33.7	289	2	JU0393
9	323	33.4	527	2	S32430
10	310.5	32.1	251	2	C39261
11	303.5	31.4	528	1	TLZSA
12	300.5	31.1	278	2	S23519
13	293.5	30.4	250	2	JN0108
14	274.5	28.4	277	2	S22494
15	273	28.3	286	1	RLPUGG
16	273	28.3	570	2	S62627
17	272	28.2	286	2	S25560
18	270	28.0	245	2	JC4840
19	265	27.4	286	2	JC4235
20	264	27.3	254	2	PD0018
21	243.5	25.2	316	2	JT0753
22	186.5	19.3	294	2	S28421
23	174	18.0	278	2	A39817
24	170	17.6	313	2	S17757
25	168.5	17.4	261	2	JED401
26	148.5	15.4	289	2	J12573
27	138	14.3	272	2	JC4811
28	131	13.6	253	2	S28542
29	127	13.1	253	2	S28539

30	127	13.1	253	2	S29931	RNA N-glycosidase
31	124	12.8	283	2	S05205	RNA N-glycosidase
32	123	12.7	253	2	S28541	RNA N-glycosidase
33	121	12.5	292	1	RLDH62	RNA N-glycosidase
34	121	12.5	310	2	S46239	ribosome-inactivat
35	116.5	12.1	293	2	S17519	RNA N-glycosidase
36	116	12.0	275	2	S33631	tritin - wheat
37	115	11.9	253	2	A58923	RNA N-glycosidase
38	114	11.8	280	1	RLDH	RNA N-glycosidase
39	112	11.6	280	2	JC5848	protein synthesis
40	110	11.4	236	2	S17932	hypothetical prote
41	109	11.3	319	2	S21940	shiga-like toxin I
42	108	11.2	319	2	I54695	shiga toxin 2 subu
43	108	11.2	319	2	E90779	Shiga toxin 2 subu
44	108	11.2	319	2	G85640	Shiga toxin 2 subu
45	107.5	11.1	318	2	S01032	Shiga-like toxin I

## ALIGNMENTS

## RESULT 1

RLCSD  
ricin D precursor - castor bean  
N:Contains: RNA N-glycosidase (EC 3.2.2.22)  
C:Species: Ricinus communis (castor bean)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
R:Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903  
R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.F.  
Nucleic Acids Res. 13, 8019-8033, 1985  
A:Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.  
A:Reference number: A24041; MUID:86067214; PMID:2999712  
A:Accession: A24041  
A:Molecule type: DNA  
A:Residues: 1-576 <HAL>  
A:Cross-references: GB:X03179; NID:g21082; PIDN:CAA26939.1; PID:g21083  
R:Tregear, J.W.; Roberts, L.M.  
Plant Mol. Biol. 18, 515-525, 1992  
A:Title: The lectin gene family of Ricinus communis: cloning of a functional ricin ge  
A:Reference number: S20513; MUID:92163016; PMID:1371405  
A:Accession: S20513  
A:Molecule type: DNA  
A:Residues: 1-576 <TRE>  
A:Cross-references: EMBL:X52908; NID:g21084; PIDN:CAA37095.1; PID:g21085  
R:Lamb, F.I.; Roberts, L.M.; Lord, J.M.  
Eur. J. Biochem. 148, 265-270, 1985  
A:Title: Nucleotide sequence of cloned cDNA coding for preproricin.  
A:Reference number: A24614; MUID:85179479; PMID:3838723  
A:Accession: A24614  
A:Molecule type: mRNA  
A:Residues: 12-75, 'D', 77-550, 'R', 552-576 <LAM>  
A:Cross-references: GB:X02388; NID:g21077; PIDN:CAA26230.1; PID:g21078  
R:Yoshitake, S.; Funatsu, G.; Funatsu, M.  
Agric. Biol. Chem. 42, 1267-1274, 1978  
A:Title: Isolation and sequences of peptic peptides, and the complete sequence of the  
A:Reference number: A03372  
A:Accession: A03372  
A:Molecule type: protein  
A:Residues: 36-97, 'Q', 99-109, 'S', 111-269, 'D', 277-283, 'L', 285-288, 290-302 <YOS>  
R:Arai, T.; Funatsu, G.  
FEBS Lett. 191, 121-124, 1985  
A:Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptop  
A:Reference number: A24010  
A:Accession: A24010  
A:Molecule type: protein  
A:Residues: 315-383, 'PS', 386-576 <ARA>  
R:Funatsu, G.; Kimura, M.; Funatsu, M.  
Agric. Biol. Chem. 43, 2221-2224, 1979  
A:Title: Primary structure of Ala chain of ricin D.  
A:Reference number: A03374  
A:Accession: A03374  
A:Molecule type: protein

A:Residues: 315-335, 'N', 337-342, 'NH', 345-362, 364-383, 'PS', 386-399, 'T', 401, 'D', 403, 'E', 405, 'V', 527, 'E', 529-564, 'W', 566, 'H', 567-570, 'LI', 573-574, 'F' <PUN>  
 A:Note: this paper, one of a series, summarizes the experimental details for the determination of the mechanism of ricin A-chain and implications for the mechanism of ricin B-chain.  
 R:Ready, M.P.; Kim, Y.; Robertus, J.D.  
 A:Title: Site-directed mutagenesis of ricin A-chain and implications for the mechanism of ricin B-chain.  
 A:Reference number: A48237; MUID:91352006; PMID:1881883  
 A:Contents: annotation: active site  
 A:Contents: annotation: active site  
 R:Robertus, J.D.  
 A:Title: Structure of ricin B-chain at 2.5 angstrom resolution.  
 A:Reference number: A48238; MUID:91352005; PMID:1881882  
 A:Contents: annotation: X-ray crystallography, 2.5 angstroms  
 R:Katzin, B.J.; Collins, E.J.; Robertus, J.D.  
 A:Title: Structure of ricin A-chain at 2.5 angstroms.  
 A:Reference number: A48239; MUID:91352004; PMID:1881881  
 A:Contents: annotation: X-ray crystallography, 2.5 angstroms  
 C:Comment: The functional molecule is a disulfide-linked dimer of A and B chains, which C:Comment: The A chain inhibits protein synthesis; it inactivates the 60S ribosomal subunit into the cell of the A chain; B chains are also responsible for cell agglutination (lectin C:Comment: This protein is cytotoxic and very poisonous to animals.  
 C:Superfamily: ricin: RNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed  
 F:1-35/Domain: signal sequence #status predicted <SIG>  
 F:36-302/Product: ricin D chain A #status experimental <ACH>  
 F:46-293/Domain: RNA N-glycosidase homology <RNG>  
 F:315-376/Product: ricin B chain B #status experimental <BCH>  
 F:331-373, 374-414, 417-455, 462-497, 501-540, 543-576/Region: 40-residue repeats  
 F:449, 409, 449/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:115, 158, 243, 244/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:212/Active site: Glu #status experimental  
 F:215/Active site: Arg #status predicted  
 F:224-318, 334-353, 377-394, 465-478, 504-521/Disulfide bonds: #status experimental  
 F:336, 349, 360/Binding site: N-acetylgalactosamine (Asp, Glu, Asn) #status experimental  
 F:548, 569/Binding site: N-acetylgalactosamine (Asp, Asn) #status experimental

Query Match 98.4%; Score 951; DB 1; Length 576;  
 Best Local Similarity 94.9%; Pred. No. 2, 5e-79;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFPKOYPIINFTTAGATVOSTNTNIRAVRGRLT-----VLPRVGLPINORFIV 50  
 DB 36 IFPKOYPIINFTTAGATVOSTNTNIRAVRGRLTGGADVREHETPLPRVGLPISORFIV 95  
 QY 51 ELSNHAELSVTLALDVTNAVYVGRAGNSAFFPHDNOEDAEATHLFTDVONRYTFAFG 110  
 DB 96 ELSNHAELSVTLALDVTNAVYVGRAGNSAFFPHDNOEDAEATHLFTDVONRYTFAFG 155  
 QY 111 GNDRLLEQLAGNLRNLELNGPLLEAIALYVYSTGCTOLPTLARSFIICMISEAR 170  
 DB 156 GNDRLLEQLAGNLRNLELNGPLLEAIALYVYSTGCTOLPTLARSFIICMISEAR 215  
 QY 171 FOYIEGEMRTIRIRNRRS 188  
 DB 216 FOYIEGEMRTIRIRNRRS 233

RESULT 2  
 RICSAG  
 agglutinin precursor - castor bean  
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C:Species: Ricinus communis (castor bean)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
 C:Accession: A24261; A24210  
 R:Robertus, J.D.; Lamb, J.F.; Pappin, D.J.C.; Lord, J.M.  
 J: Biol. Chem. 260, 15682-15686, 1985  
 A:Title: The primary sequence of Ricinus communis agglutinin. Comparison with ricin.  
 A:Reference number: A24261; MUID:8605949; PMID:2999130  
 A:Accession: A24261  
 A:Molecule type: rRNA  
 A:Residues: 1-564 <ROB>  
 A:Cross-references: GB:M12089; NID:g169700; PIDN:AAA33869.1; PID:g169701

R:Araki, T.; Yoshitake, Y.; Funatsu, G.  
 Biochim. Biophys. Acta 872, 277-285, 1986  
 A:Title: The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin.  
 A:Reference number: A24210  
 A:Accession: A24210  
 A:Molecule type: Protein  
 A:Residues: 303-325, 'F', 327-330, 'T', 332-361, 'D', 363-373, 'G', 375-403, 'T', 405-551, 'V', 5  
 A:Comment: This protein has strong agglutinating activity and weak cytotoxicity compa  
 C:Superfamily: ricin: RNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; s  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-290/Product: agglutinin chain A #status predicted <ACH>  
 F:35-281/Domain: RNA N-glycosidase homology <RNG>  
 F:303-564/Product: agglutinin chain B #status experimental <BCH>  
 F:319-361, 362-402, 403-443, 450-485, 489-528, 531-564/Region: 40-residue repeats  
 F:34, 259/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:104, 147, 231, 232/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:200, 203/Active site: Glu, Arg #status predicted  
 F:282-306, 322-341, 365-382, 453-466, 492-509/Disulfide bonds: #status predicted  
 F:324, 337, 348/Binding site: N-acetylgalactosamine (Asp, Glu, Asn) #status predicted  
 F:397, 437/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:536, 557/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 89.2%; Score 861.5; DB 1; Length 564;  
 Best Local Similarity 86.9%; Pred. No. 3, 9e-71;  
 Matches 172; Conservative 7; Mismatches 8; Indels 11; Gaps 2;

QY 1 IFPKOYPIINFTTAGATVOSTNTNIRAVRGRLT-----VLPRVGLPINORFIV 50  
 DB 25 IFPKOYPIINFTTAGATVOSTNTNIRAVRGRLTGGADVREHETPLPRVGLPISORFIV 84  
 QY 51 ELSNHAELSVTLALDVTNAVYVGRAGNSAFFPHDNOEDAEATHLFTDVONRYTFAFG 110  
 DB 85 ELSNHAELSVTLALDVTNAVYVGRAGNSAFFPHDNOEDAEATHLFTDVONRYTFAFG 144  
 QY 111 GNDRLLEQLAGNLRNLELNGPLLEAIALYVYSTGCTOLPTLARSFIICMISEAR 170  
 DB 145 GNDRLLEQLAGNLRNLELNGPLLEAIALYVYSTGCTOLPTLARSFIICMISEAR 203  
 QY 171 FOYIEGEMRTIRIRNRRS 188  
 DB 204 FOYIEGEMRTIRIRNRRS 221

RESULT 3  
 RLRT7T  
 rRNA N-glycosidase (EC 3.2.2.22) alpha-trichosanthin precursor [validated] - Mongolia  
 N:Alternate names: alpha-TCs; type I ribosome-inactivating protein  
 C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)  
 C:Date: 30-Sep-1988 #sequence\_revision 26-Jan-1996 #text\_change 23-Mar-2001  
 C:Accession: J70566; A36274; JCI093; A36273; J70003  
 R:Shaw, P.C.; Yung, M.H.; Zhu, R.H.; Ho, W.K.K.; Ng, T.B.; Yeung, H.W.  
 Gene 97, 267-272, 1991  
 A:Title: Cloning of trichosanthin cDNA and its expression in Escherichia coli.  
 A:Reference number: J70566; MUID:91153657; PMID:1999291  
 A:Accession: J70566  
 A:Molecule type: mRNA  
 A:Residues: 1-289 <SHA>  
 A:Cross-references: GB:M34858; NID:g170536; PIDN:AAA34207.1; PID:g170537  
 A:Experimental source: tuber  
 R:Chow, T.P.; Feldman, R.A.; Lovett, M.; Platek, M.  
 J: Biol. Chem. 265, 8670-8674, 1990  
 A:Title: Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type I  
 A:Reference number: A36274; MUID:90256790; PMID:2341400  
 A:Accession: A36274  
 A:Molecule type: DNA  
 A:Residues: 1-233, 'T', 235-246, 'W', 248-289 <CHO>  
 A:Cross-references: GB:J05434; NID:g170534; PIDN:AAA34206.1; PID:g170535  
 R:zheng, H.G.; Wang, B.; Shao, P.Z.; Yang, X.R.  
 Acta Genet. Sin. 21, 42-51, 1994  
 A:Title: Cloning and DNA sequencing of the gene encoding Trichosanthin.  
 A:Reference number: JCI093; MUID:94271613; PMID:8003348  
 A:Accession: JCI093

C:Species: Abrus precatorius (Indian licorice)  
C:Date: 30-Sep-1993 #sequence\_revision 01-Aug-1997 #text\_change 01-Aug-1997  
C:Accession: S32431, S34408  
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
J. Mol. Biol. 229, 263-267, 1993  
A:Title: Primary structure of three distinct isobrans determined by cDNA sequencing  
A:Reference number: S32429; MUID:93132798; PMID:8421313  
A:Accession: S32431  
A:Molecule type: mRNA  
A:Residues: 1-528 <HUN>  
R:Hung, C.; Lee, M.; Lee, T.; Lin, J.  
submitted to the EMBL Data Library, March 1993  
A:Reference number: S34408  
A:Accession: S34408  
A:Molecule type: mRNA  
A:Residues: 1-169, 'C', 171-320, 'U', 322-528 <HUT>  
A:Cross-references: GB:M8346  
C:Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating the A and B chains are linked by a single disulfide bond, which is essential for toxicity.  
C:Superfamily: ricin; RNA N-glycosidase homology  
C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectinase  
F:1-251/Product: abrin-d chain A #status predicted <ACH>  
F:7-246/Domain: RNA N-glycosidase homology <RNG>  
F:261-528/Product: abrin-d chain B #status predicted <BCH>  
F:283-325, 326-366, 369-407, 414-449, 453-492, 495-528/Region: 40-residue repeats  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted  
F:164, 167/Active site: Glu, Arg #status predicted  
F:174, 113, 195, 196/binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
F:200, 223, 361, 401, 402/binding site: carbonylate (Asn) (covalent) #status predicted  
F:247-269, 286-305, 329-346, 417-430, 456-473/Disulfide bonds: #status predicted  
F:288, 312/binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
F:500, 521/binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 33.9%; Score 327.5; DB: 2; Length 528;  
Best Local Similarity 43.9%; Pred. No. 3.5e-22;  
Matches 82; Conservative 22; Mismatches 68; Indels 15; Gaps 4;

QY 5 QPPIINFTTGAATVQSYNTFNRAVRGRLT-----VLPNRVGLPINORFLVELSNHA 56  
Db 1 QDVYIKRTTBATTSQSKQFLEALRQRLTGLHDIVLPDPTVEERNKRTVELNSSE 60  
QY 57 ELVYTLADVTNNAVYGVYRAGNSAYFFHPNQEDAEATITHELPDVQNKRYTFAFGNDR 116  
Db 61 RESIEVGIDVTNAVYVAVRAGSQSYFL---RDAPASASTYLFPEQTQ-RYSLRFDGSYGLD 116  
QY 117 EQLAGNIRENTELGNGPLPEALISALYYSTGCTGLOPLTARSFFICIQMISNAAFQYIEG 176  
Db 117 ERMHQTRREELSLGLALTAHLS--FLRSGASNDEKARPLIVIIWMASEARARYRISN 173  
QY 177 EMRTTRIR 183  
Db 174 RVGVYSIR 180

RESULT 5  
S16022  
abrin-c precursor - Indian licorice  
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
C:Species: Abrus precatorius (Indian licorice)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Aug-1999  
C:Accession: S16022  
R:Wood, K.A.; Lord, J.M.; Mawzynczak, E.J.; Platak, M.  
Eur. J. Biochem. 198, 723-732, 1991  
A:Title: Preproabrin: genomic cloning, characterisation and the expression of the A-c  
A:Reference number: S16022; MUID:91266957; PMID:2050149  
A:Accession: S16022  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-562 <MOO>  
A:Cross-references: EMBL:X55667; MID:g16084; PIDN:CA39202.1. PID:g16085  
C:Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating the A and B chains are linked by a single disulfide bond, which is essential for toxicity.

C:Superfamily: r1cin; rRNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglytamic acid  
 F:35-285/Product: abrin-c chain A #status predicted <ACH>  
 F:4-280/Domain: rRNA N-glycosidase homology <RNG>  
 F:295-562/Product: abrin-c chain B #status predicted <BCH>  
 F:317-359,360-400,403-441,448-483,487-526,529-562/Region: 40-residue repeats  
 F:35/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
 F:108,147,229,230/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:198,201/Active site: Glu, Arg #status predicted  
 F:224,287,395,435,436/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:281-303,320-339,363-380,451-464,490-507/Disulfide bonds: #status predicted  
 F:322,346/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 F:534,555/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 33.7%; Score 327.5; DB 2; Length 562;  
 Best Local Similarity 43.9%; Pred. No. 3.8e-22;  
 Matches 82; Conservative 22; Mismatches 68; Indels 15; Gaps 4;

OY 5 QYPIINFTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINOFLVELSNHA 56  
 DB 35 QDQYIKFTTSGATSSQSKQFLEALRQLTGGLIHDPDPPTVEERNRYITVELSNSE 94  
 OY 57 ELSTYTLADVTNAYVGYRAGNSAYFFHPNOEDAETITLFDVQNRITFAFGNYDL 116  
 DB 95 RESIEVGIDVTNAYVAVRAGSOSYFL--RDAPASATYLFPGTG-RTSLRFDGSYGL 150  
 OY 117 EQLAGNREINELGNGLEAFAISALYYSTGTOLPTLARSFTICOMISEARFYIEG 176  
 DB 151 ERMAHGTREITSIAGLQALTHAIS---FLRSGASNDKAKTLIVITOMSEARFYISN 207  
 OY 177 EMRTIR 183  
 DB 208 RVGSIR 214

## RESULT 6

JU0393  
 karasurin - Mongolian snake-gourd  
 C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)  
 C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 12-Apr-1995  
 C:Accession: JU0393; PS0163  
 R:Toyokawa, S.; Takeda, T.; Kato, Y.; Wakabayashi, K.; Ogihara, Y.  
 Chem. Pharm. Bull. 39, 1244-1249, 1991  
 A:Title: The complete amino acid sequence of an abortifacient protein, karasurin.  
 A:Reference number: JU0393; MUID:92005921; PMID:1914000  
 A:Accession: JU0393  
 A:Molecule type: protein  
 A:Residues: 1-247 <TOY>  
 A:Note: a sequence which lacks Ala-247 is also shown in this publication  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 C:Keywords: abortifacient  
 F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 33.7%; Score 326; DB 2; Length 247;  
 Best Local Similarity 39.6%; Pred. No. 1.8e-22;  
 Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

OY 9 INFTAGATVQSYTNFIRAVR-----GRLTVLP-NRVGLPINOFLVELSNHAELSVT 61  
 DB 2 VSFRLSGATSSQSKQFLEALRQLTGGLIHDPDPPTVEERNRYITVELSNSE 61  
 DB 62 VAIQVTNAYVGYRAGNSAYFFHPNOEDA-ETITLFTDVQNRITFAFGNYDLRQLA 120  
 OY 121 GNLREINELGNGLEAFAISALYYSTGTOLPTLARSFTICOMISEARFYIEGEMRT 180  
 DB 119 GKIRENIPGLPALDSAITTLFTYNNAN-----SAASALMWLIQSTSPAAKRFLEOIGK 173

OY 181 RI 182  
 DB 174 RV 175

## RESULT 7

JC5032  
 karasurin-B - Trichosanthes kirilowii var. japonica  
 C:Species: Trichosanthes kirilowii var. japonica  
 C:Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 23-May-1997  
 C:Accession: JC5032  
 R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.  
 Biol. Pharm. Bull. 19, 1485-1489, 1996  
 A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and  
 A:Reference number: JC5032; MUID:97108848; PMID:8951169  
 A:Accession: JC5032  
 A:Molecule type: protein  
 A:Status: preliminary  
 A:Residues: 1-247 <KON>  
 C:Comment: This protein belongs to type I ribosomal-inactivating proteins which catal  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 33.7%; Score 326; DB 2; Length 247;  
 Best Local Similarity 39.6%; Pred. No. 1.8e-22;  
 Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

OY 9 INFTAGATVQSYTNFIRAVR-----GRLTVLP-NRVGLPINOFLVELSNHAELSVT 61  
 DB 2 VSFRLSGATSSQSKQFLEALRQLTGGLIHDPDPPTVEERNRYITVELSNSE 61  
 OY 62 IALDVTNAYVGYRAGNSAYFFHPNOEDA-ETITLFTDVQNRITFAFGNYDLRQLA 120  
 DB 62 VAIQVTNAYVGYRAGNSAYFFHPNOEDA-ETITLFTDVQNRITFAFGNYDLRQLA 118  
 OY 121 GNLREINELGNGLEAFAISALYYSTGTOLPTLARSFTICOMISEARFYIEGEMRT 180  
 DB 119 GKIRENIPGLPALDSAITTLFTYNNAN-----SAASALMWLIQSTSPAAKRFLEOIGK 173  
 OY 181 RI 182  
 DB 174 RV 175

## RESULT 8

JC5606  
 karasurin C - Trichosanthes kirilowii var. japonica  
 N:Contains: karasurin A  
 C:Species: Trichosanthes kirilowii var. japonica  
 C:Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 19-Jul-2002  
 C:Accession: JC5606; JC5033  
 R:Mizukami, H.; Kondo, T.; Ogihara, Y.  
 Biol. Pharm. Bull. 20, 711-713, 1997  
 A:Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating pr  
 A:Reference number: JC5606; MUID:97356562; PMID:9212998  
 A:Accession: JC5606  
 A:Molecule type: DNA  
 A:Residues: 1-289 <MIZ>

A:Cross-references: DDBJ:AB000666; NID:92329830; PINO:BA21786.1; PID:92329831  
 R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.  
 Biol. Pharm. Bull. 19, 1485-1489, 1996  
 A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and  
 A:Reference number: JC5032; MUID:97108848; PMID:8951169  
 A:Accession: JC5033  
 A:Molecule type: preliminary  
 A:Residues: 22-270 <KON>  
 C:Comment: This protein is a ribosome-inactivating protein and exhibit cytotoxic, abo  
 C:Comment: This protein belongs to type I ribosomal-inactivating proteins which catal  
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
 F:22-270/Product: karasurin C #status predicted <MAC>  
 F:24-270/Product: karasurin A #status predicted <MAA>  
 F:27-266/Domain: rRNA N-glycosidase homology <RNG>

Query Match 33.7%; Score 326; DB 2; Length 289;  
 Best Local Similarity 39.6%; Pred. No. 2.2e-22;  
 Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

QY 9 INFTAGATVOSTYNTFIRAVR-----GRLTVLP-NRVLPIPNOREFLVELSNHAEISVT 61  
 Db 25 VSRFLSGATSSSYGVFTISNLRKALPERKLTLDIPLRLSTLPSORALHLINVADETIS 84  
 QY 62 LADVTNAYVVGVRAGNSAFAFFHPDNOEDA-EAITHLFTDVONKRTFAFGNDRLLEOLA 120  
 Db 85 VALDVNAYVVGVRAGNSAFAFFHPDNOEDA-EAITHLFTDVONKRTFAFGNDRLLEOLA 141  
 QY 121 GNLRNIEIENGPELEAISALYYSTGQTLPLARSPFIQIOMISEARFOYIGEMRT 180  
 Db 142 GKRENIPIGLPLALDSAITTLFTYYNAN-----SAASALWVLQISTSEARVFIQOICK 196  
 QY 181 RI 182  
 Db 197 RV 198

## RESULT 9

S32430  
 abrin-b precursor - Indian licorice (fragment)  
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C:Species: Abrus precatorius (Indian licorice)  
 C>Date: 30-Sep-1993 #sequence\_revision 01-Aug-1997 #text\_change 20-Aug-1999  
 C:Accession: S32430; JCI399  
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
 J. Mol. Biol. 229, 263-267, 1993  
 A>Title: Primary structure of three distinct isoabirins determined by cDNA sequencing. CC  
 A:Reference number: S32429; MUID:93132798; PMID:8421313  
 A:Accession: S32430  
 A:Molecule type: mRNA  
 A:Residues: 1-527 <HUN>  
 A:Cross-references: GB:M98345; NID:9166296; PIDN:AAA32625.1; PID:9166297  
 R:Kimura, M.; Sumizawa, T.; Funatsu, G.  
 Biosci. Biotechnol. Biochem. 57, 166-169, 1993  
 A>Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic  
 A:Reference number: JCI398; MUID:93169023; PMID:7763422  
 A:Accession: JCI399  
 A:Molecule type: protein  
 A:Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'W', 427, 'D', 429-430  
 A:Experimental source: seed  
 C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; F  
 F:1-250/Product: abrin-b chain A #status predicted <ACH>  
 F:1-249/Domain: rRNA N-glycosidase homology <RNG>  
 F:260-527/Product: abrin-b chain B #status experimental <BCH>  
 F:282-324,325-365,368-406,413-448,452-491,494-527/Region: 40-residue repeats  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted  
 F:74,113,194,195/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:110,360,400/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:163,166/Active site: Glu, Arg #status predicted  
 F:246-268,285-304,328-345,416-429,455-472/Disulfide bonds: #status predicted  
 F:287,311/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted  
 F:499,520/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

## Query Match

Best Local Similarity 33.4%; Score 323; DB 2; Length 527;  
 Matches 83; Conservative 21; Mismatches 69; Indels 16; Gaps 4;

QY 5 OYPIINFTAGATVOSTYNTFIRAVGRLL-----VLPNRVGLPIPNOREFLVELSNHAEISVT 56  
 Db 1 QOVNLIKFTTEGATSSSYGVFTISNLRKALPERKLTLDIPLRLSTLPSORALHLINVADETIS 60  
 QY 57 ELSTLADVTNAYVVGVRAGNSAFAFFHPDNOEDA-EAITHLFTDVONKRTFAFGNDRLLEOLA 116  
 Db 61 TESIENGIDVSNAYVVGVRAGNSAFAFFHPDNOEDA-EAITHLFTDVONKRTFAFGNDRLLEOLA 116  
 QY 117 EOLAGLRNIEIENGPELEAISALYYSTGQTLPLARSPFIQIOMISEARFOYIGEMRT 176  
 Db 117 ERLLARTRQOIPGLALRHAISFL-----QSGTDOELARLIVLIOMASERARFISY 172  
 QY 177 EMKTRIRYN 185

Db 173 RVGVSIRTN 181

## RESULT 10

C39761  
 abrin (clone 7.2) precursor - Indian licorice (fragment)  
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C:Species: Abrus precatorius (Indian licorice)  
 C>Date: 21-Feb-1992 #sequence\_revision 31-Dec-1993 #text\_change 20-Aug-1999  
 C:Accession: C39761; S14471  
 R:Evensen, G.; Mathiesen, A.; Sundan, A.  
 J. Biol. Chem. 266, 6848-6852, 1991  
 A>Title: Direct molecular cloning and expression of two distinct abrin A-chains.  
 A:Reference number: A39761; MUID:91201329; PMID:2016300  
 A:Accession: C39761  
 A:Molecule type: DNA  
 A:Residues: 1-251 <EVE>  
 R:Evensen, G.; Mathiesen, A.; Sundan, A.  
 submitted to the EMBL Data Library, October 1990  
 A:Description: Direct molecular cloning of two distinct abrin A-chains.  
 A:Reference number: S14471  
 A:Accession: S14471  
 A:Molecule type: DNA  
 A:Residues: 'W', 1-251 <EVE>  
 A:Cross-references: EMBL:X54872; NID:916088; PIDN:CAA38654.1; PID:916089  
 C:Superfamily: ricin; rRNA N-glycosidase homology  
 C:Keywords: duplication; glycosidase; hydrolase; lectin; toxin  
 F:1-251/Product: abrin (clone 7.2) chain A #status predicted <ACH>  
 F:7-246/Domain: rRNA N-glycosidase homology <RNG>  
 F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:164,167/Active site: Glu, Arg #status predicted

## Query Match

Best Local Similarity 32.1%; Score 310; DB 2; Length 251;  
 Matches 79; Conservative 22; Mismatches 67; Indels 15; Gaps 4;

QY 9 INFTAGATVOSTYNTFIRAVGRLL-----VLPNRVGLPIPNOREFLVELSNHAEISVT 60  
 Db 5 IKFSTEGATSSSYGVFTISNLRKALPERKLTLDIPLRLSTLPSORALHLINVADETIS 64  
 QY 61 TLADVTNAYVVGVRAGNSAFAFFHPDNOEDA-EAITHLFTDVONKRTFAFGNDRLLEOLA 120  
 Db 65 EVGIDVTNAYVVGVRAGNSAFAFFHPDNOEDA-EAITHLFTDVONKRTFAFGNDRLLEOLA 120  
 QY 121 GNLRNIEIENGPELEAISALYYSTGQTLPLARSPFIQIOMISEARFOYIGEMRT 180  
 Db 121 HGTNRQISLGLALRHAISFL-----FLRSGASNDEEKARLIVLIOMASERARFISY 177  
 QY 181 RIR 183  
 Db 178 SIR 180

## RESULT 11

TLISA  
 abrin-a precursor - Indian licorice (fragment)  
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C:Species: Abrus precatorius (Indian licorice)  
 C>Date: 31-Dec-1993 #sequence\_revision 01-Aug-1997 #text\_change 16-Jul-1999  
 C:Accession: S32429; JCI398; S14472; S24133; S74110; S74111  
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
 J. Mol. Biol. 229, 263-267, 1993  
 A>Title: Primary structure of three distinct isoabirins determined by cDNA sequencing.  
 A:Reference number: S32429; MUID:93132798; PMID:8421313  
 A:Accession: S32429  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 'E', 2-528 <HUN>  
 A:Cross-references: GB:M98344; NID:9166294; PIDN:AAA32624.1; PID:9166295  
 A>Note: the coding region for the sequence shown is preceded by an ATG codon  
 R:Funatsu, G.; Taguchi, Y.; Kamenosono, M.; Yanaka, M.  
 Agric. Biol. Chem. 52, 1095-1097, 1988





Best Local Similarity 33.7%; Pred. No. 1,8e-19;  
Matches 63; Conservative 46; Mismatches 57; Indels 21; Gaps 4;

QY 9 INFTTGATVQSYTNFIRAVRGRL-----TVLPNRVGLPIQRFITVELSNHAE 58  
Db 3 VFSLSGASQSKSYKRTITLAKRLPSEKVENYSPILLPFSASG---SRITLMQLSYDAK 59  
QY 59 SVTLALDVNNAVYVGYRAGNSAFPHFDNOEDAEATHTLFTDVONRYTFAGGNDRLERQ 118  
Db 60 AITMADVNTVNYIMGVNVTSTSYF---ANESDAKLASQYVFKSGSLVTIPYSGNERLQN 116  
QY 119 LAGNLEPNTLELGPLEPAISALYYXSTGQTLPILARSPITICQIMISAPQYIEGM 178  
Db 117 AAGKIREKIPLEGFRALDSALTIFHYDS-----TAAAFVLTQTTEASRKYIEGOI 171  
QY 179 RTIRIRYN 185  
Db 172 IERIRPKN 178

## RESULT 14

RNA N-glycosidase (EC 3.2.2.22) alpha-luffin precursor - smooth loofah  
N:Alternate names: protein synthesis inhibitor; ribosome-inactivating protein luffin-A  
C:Species: Luffa cylindrica (smooth loofah)  
C>Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 20-Aug-1999  
C:Accession: S22494; S26390; J00202; A32542  
R:Atoko, J.; Habuka, N.; Miyano, M.; Masuta, C.; Kotwal, A.  
Plant Mol. Biol. 18, 1199-1202, 1992  
A:Title: Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-inactivating prot  
A:Reference number: S22494; MUID:92288316; PMID:1600156  
A:Accession: S22494

A:Molecule type: mRNA  
A:Residues: 1-277 <MAT>  
A:Cross-references: EMBL:X62371; NID:g19145; PIDN:CAA44229.1; PID:g19146  
R:Islam, M.R.; Nishida, H.; Funatsu, G.  
Agric. Biol. Chem. 54, 2967-2978, 1990  
A:Title: Complete amino acid sequence of luffin-a, a ribosome-inactivating protein from  
A:Reference number: S26390; MUID:91248471; PMID:1368651  
A:Accession: S26390  
A:Molecule type: protein  
A:Residues: 21-53, 'L', '55', 'I', '57-86', 'SQL', '90-154', 'L', '156-157', 'I', '159-173', 'L', '175-209', 'SL'  
R:Islam, M.R.; Nishida, H.; Funatsu, G.  
Agric. Biol. Chem. 54, 1343-1345, 1990  
A:Title: Complete amino acid sequence of luffin-a, a ribosome-inactivating protein from  
A:Reference number: J00202; MUID:91197482; PMID:1368623  
A:Accession: J00202  
A:Molecule type: protein  
A:Residues: 21-53, 'L', '55', 'I', '57-86', 'SQL', '90-154', 'L', '156-157', 'I', '159-173', 'L', '175-209', 'SL'  
R:Ramakrishnan, S.; Enghild, J.J.; Bryant Jr., H.L.; Xu, F.J.  
Biochem. Biophys. Res. Commun. 160, 509-516, 1989  
A:Title: Characterization of a translation inhibitory protein from Luffa aegyptiaca.  
A:Reference number: A32542; MUID:89246493; PMID:2719679  
A:Accession: A32542  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 21-28, 'G', '30', 'X', '32-34', 'K', '36-40' <RAM>  
C:Function:  
A:Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA ther  
C:Superfamily: RNA N-glycosidase: RNA N-glycosidase homology  
C:Keywords: glycoprotein; glycosidase; hydrolase; seed; toxin  
F:1-20/Domain: signal sequence #status: predicted <SIG>  
F:21-267/Product: RNA N-glycosidase alpha-luffin #status experimental <MAT>  
F:24-262/Domain: RNA N-glycosidase homology <RNS>  
F:268-277/Domain: carboxyl-terminal propeptide #status predicted <CPT>  
F:48-53, '7', '104-225, 246/Binding site: carbohydrate #asn (covalent) #status experimental  
F:90, 179, 182/Active site: Tyr, Glu, Arg #status predicted

Query Match 28.4%; Score 274.5; DB 2; Length 277;  
Best Local Similarity 33.3%; Pred. No. 1.1e-17;  
Matches 63; Conservative 42; Mismatches 61; Indels 23; Gaps 5;

QY 9 INFTTGATVQSYTNFIRAVRGRLTVLPNRVGLPIN-----QRFITVELSNHAE 57  
Db 22 VFSLSGSSSTSYSKFIIDLR---KALPSN-GIYVNTILLSSASGSRITLMLSNYDG 77  
QY 58 LSVTLALDVNNAVYVGYRAGNSAFPHFDNOEDAEATHTLFTDVONRYTFAGGNDRLERQ 117  
Db 78 KAIPVADVNTVNYIMGVNVTSTSYF---NESDAKLASQYVFKSGSLVTIPYSGNERLQ 134  
QY 118 OLAGNLEPNTLELGPLEPAISALYYXSTGQTLPILARSPITICQIMISAPQYIEGE 177  
Db 135 TAAGKIREKIPLEGFRALDSALTIFHYDS-----TAAAFVLTQTTEASRKYIEGOI 189  
QY 178 MTRIRIRYN 186  
Db 190 IERIRSKN 198

## RESULT 15

RNA N-glycosidase (EC 3.2.2.22) alpha-momorcharin precursor [validated] - balsam pea  
N:Alternate names: agglutinin; momordin-A; ribosome-inactivating protein momorcharin  
C:Species: Momordica charantia (balsam pear, bitter melon)  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 15-Sep-2000  
C:Accession: S14273; A61318; S16490; J00628; S01670  
R:Ho, W.K.K.; Liu, S.C.; Shaw, P.C.; Yeung, H.W.; Ng, T.B.; Chan, W.Y.  
Biochim. Biophys. Acta 1088, 311-314, 1991  
A:Title: Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating protein.  
A:Reference number: S14273; MUID:91159486; PMID:2001404  
A:Accession: S14273

A:Molecule type: mRNA  
A:Residues: 1-286 <HOM>  
A:Cross-references: EMBL:X57682; NID:g19527; PIDN:CAA40869.1; PID:g19528  
R:Li, S.S.L.  
Experientia 36, 524-527, 1980  
A:Title: Purification and partial characterization of two lectins from Momordica char  
A:Reference number: A61318; MUID:80201763; PMID:7379938  
A:Accession: A61318  
A:Molecule type: protein  
A:Residues: 24-50 <LIR>  
A>Note: as a lectin shows agglutinating activity for type-O red blood cells  
R:Montecucchi, P.C.; Lazzarini, A.M.; Barbieri, L.; Stille, F.; Soria, M.; Leppl, D.  
Int. J. Pept. Protein Res. 33, 263-267, 1989  
A:Title: N-terminal sequence of some ribosome-inactivating proteins.  
A:Reference number: S16331; MUID:89326691; PMID:2753596  
A:Accession: S16490  
A:Molecule type: protein  
A:Residues: 24-68, 'X', '70' <MON>  
R:Minami, Y.; Funatsu, G.  
Biosect. Biotechnol. Biochem. 57, 1141-1144, 1993  
A:Title: The complete amino acid sequence of momordin-a, a ribosome-inactivating prot  
A:Reference number: J00628; MUID:93372485; PMID:7763984  
A:Accession: J00628  
A:Molecule type: protein  
A:Residues: 24-107, 'Q', '109-123, 125-147', 'L', '149-154', 'I', '156-205', 'I', '207-208', 'L', '210-21  
R:Ren, J.; Wang, Y.; Dong, Y.; Stuart, D.I.  
Submitted to the Brookhaven Protein Data Bank, January 1994  
A:Reference number: A52272; PDB:1AHC  
R:Husain, J.; Tickle, I.J.; Wood, S.P.  
Submitted to the Brookhaven Protein Data Bank, March 1994  
A:Reference number: A52385; PDB:1K0M  
A:Contents: annotation; X-ray crystallography, 2.16 angstroms, residues 24-86, 'L', '88-  
R:Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.  
Submitted to the Brookhaven Protein Data Bank, July 1994  
A:Reference number: A67099; PDB:1MKH  
A:Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 24-77, 'R', '79-1  
C:Function:  
A:Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA t  
C:Superfamily: RNA N-glycosidase: RNA N-glycosidase homology  
C:Keywords: glycoprotein; glycosidase; hydrolase; lectin; seed; toxin  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-23/Domain: amino-terminal propeptide #status predicted <CPT>

F:24-269/Product: rRNA N-glycosidase alpha-momorcharin #status experimental <MAT>  
 F:27-266/Domain: rRNA N-glycosidase homology <RNG>  
 F:270-286/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
 F:93,183,186/Active site: Tyr, Glu, Arg #status predicted  
 F:250/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 28.3% Score 273; DB 1; Length 286;

Best Local Similarity 34.8%; Pred. No. 1.6e-17;  
 Matches 63; Conservative 40; Mismatches 62; Indels 16; Gaps 4;

```

QY 9 INFTAGATVQSYTNFIRAVRGRLTVLPNRVGLPI-----NQRELYVELSNHAEISVT 61
   ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 25 VSFRLSGADPRSYGCMFKDLRNALPFREKYYNIPLLPSVSGAGRYLMLHMFNYDGKIT 84
   ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 62 LALDVTNAYVYGRAGNSATFFHPDNOEDAE-AITHLFTDVQNRKTFAGGNYDRLEQLA 120
   ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 85 VAVDVTVVYIMGYLADTSTYF--NEPAELASQYVFRDARRKITLPYSGNYERLQIAA 141
   ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 121 GNLRNIEIENGPLEAITSALYYSTGTQLPTLARSFICIQMISPARFOYTEGEMPT 180
   ||||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 142 GKPREKIPIGLPALDSAITLLHDS-----TAAAGALLVLIQTAAARFKYIEQIQIE 196
   ||||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 181 R 181
   |
Db 197 R 197

```

Search completed: September 16, 2003, 11:50:34  
 Job time : 13.1928 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:30:32 ; Search time 6.59071 Seconds  
(without alignments)  
1341.437 Million cell updates/sec

Title: US-10-083-336a-4

Perfect score: 966  
Sequence: 1 IFPKQYPIINFATTGATVQS.....ARFYEGEMRTIRYNRRS 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	951	98.4	576	1	P02879 ricinus com
2	861.5	89.2	564	1	P06750 ricinus com
3	336	34.8	289	1	P09989 trichosan
4	327.5	33.9	562	1	P28590 abrus pre
5	326	33.7	289	1	P24478 trichosan
6	323	33.4	527	1	P06077 abrus pre
7	320.5	33.2	282	1	P98184 bryonia dio
8	307.5	31.8	563	1	P33183 sambucus ni
9	305	31.6	290	1	P33185 bryonia dio
10	303.5	31.4	528	1	P33140 abrus pre
11	293.5	30.4	250	1	P22851 luffa cylin
12	281.5	29.1	286	1	P09f44 cucumis fig
13	274.5	28.4	277	1	P00465 luffa cylin
14	273	28.3	286	1	P16094 momordica c
15	272	28.2	286	1	P28339 momordica b
16	264	27.3	254	1	P81446 viscum albu
17	257	26.6	294	1	P56626 trichosan
18	243.5	25.2	316	1	P33186 gelonium mu
19	186.5	19.3	294	1	P03464 phytoacca
20	174	18.0	278	1	P21326 mirabilis j
21	170	17.6	313	1	P10297 phytoacca
22	168.5	17.4	261	1	P43399 phytoacca
23	131	13.6	253	1	O41389 saponaria o
24	127	13.1	253	1	P20656 saponaria o
25	124	12.8	299	1	P27555 saponaria o
26	121	12.5	292	1	P27555 saponaria o
27	121	12.5	310	1	P27555 saponaria o
28	116.5	12.1	293	1	P24476 dianthus ca
29	114	11.8	280	1	P04399 hordeum vul
30	110	11.4	236	1	P27560 saponaria o
31	108	11.2	319	1	P09385 bacterioph
32	107	11.1	280	1	P22244 hordeum vul
33	103.5	10.7	300	1	P25891 zea mays (m

34	101.5	10.5	301	1	RIPX_MAIZE	P28522 zea mays (m
35	97	10.0	304	1	RIP9_MAIZE	P25892 zea mays (m
36	94	9.7	315	1	SLTA_BPH19	P08026 bacterioph
37	93	9.6	315	1	SLTA_BPH30	P10149 bacterioph
38	85.5	8.9	1165	1	SVT_AO04E	O67411 aquilex aeo
39	81.5	8.4	305	1	Y348_MYCPN	P75255 mycoplasma
40	79	8.2	560	1	J160_HORVU	Q00531 hordeum vul
41	78	8.1	747	1	ORPB_HUMAN	Q95XB4 homo sapien
42	76.5	7.9	609	1	HAPT_VIRCH	P24153 vibrio chol
43	76	7.9	609	1	RIP4_SAPOF	P27561 saponaria o
44	75.5	7.8	609	1	NRV_VIBPR	Q00971 vibrio prot
45	75	7.8	224	1	GPH_HAEIN	P44755 haemophilus

## ALIGNMENTS

```

RESULT 1
RICI_RICCO
AC P02879; P02880; STANDARD; PRT: 576 AA.
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ricin precursor [Contains: Ricin A chain (RNA N-glycosidase)
DE (EC 3.2.2.22); Ricin B chain].
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosidales; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP MEDLINE=86067214; PubMed=2999712;
RX Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,
RA Weaver R.F.;
RT "Genomic Cloning and characterization of a ricin gene from Ricinus
RT communis.";
RL Nucleic Acids Res. 13:8019-8033(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92163016; PubMed=1371405;
RA Tregear J.W., Roberts L.M.;
RT "The lectin gene family of Ricinus communis: cloning of a functional
RT ricin gene and three lectin pseudogenes.";
RL Plant Mol. Biol. 18:515-525(1992).
RN [3]
RP SEQUENCE OF 12-576 FROM N.A.
RX MEDLINE=85179479; PubMed=3838723;
RA Lamb A., Roberts L.M., Lord J.M.;
RT "Nucleotide sequence of cloned cDNA coding for preproricin.";
RL Eur. J. Biochem. 148:265-270(1985).
RN [4]
RP SEQUENCE OF 36-302.
RA Yoshitake S., Funatsu G., Funatsu M.;
RT "Isolation and sequences of peptic peptides, and the complete
RT sequence of the chain of ricin-D.";
RL Agric. Biol. Chem. 42:1267-1274(1978).
RN [5]
RP SEQUENCE OF 315-576.
RA Funatsu G., Kimura M., Funatsu M.;
RT "Primary structure of Ala chain of ricin D.";
RL Agric. Biol. Chem. 43:2221-2224(1979).
RN [6]
RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.
RX MEDLINE=90344223; PubMed=1368517;
RA Kimura Y., Kusnoku H., Tada M., Takagi S., Funatsu G.;
RT "Structural analyses of sugar chains from ricin A-chain variant.";
RL Agric. Biol. Chem. 54:157-162(1990).
RN [7]
RP REVIEW.
RX MEDLINE=21480122; PubMed=11595634;
RA Olsnes S., Kozlov J.V.;

```



Query Match 98.4%; Score 951; DB 1; Length 576;  
 Best Local Similarity 94.9%; Pred. No. 2, 2e-80;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFFKQYPIINFTAGATVOSTNFIKAVRGRLT-----VLPNRVGLPINORITLV 50  
 |||||||  
 DB 36 IFFKQYPIINFTAGATVOSTNFIKAVRGRLTGTGADVREHETPLVLPNRVGLPINORITLV 95  
 |||||||

QY 51 ELSNHAELSVTLALDVTNAYVYGRAGNSAFFHPDNOEDAETHTLFTDVONRTYPAFG 110  
 |||||||  
 DB 96 ELSNHAELSVTLALDVTNAYVYGRAGNSAFFHPDNOEDAETHTLFTDVONRTYPAFG 155  
 |||||||

QY 111 GNYDRLEQLAGNLRENIELGNGPLEAISALYYSTGCTOLPTLARSFICQIMISEAR 170  
 |||||||  
 DB 156 GNYDRLEQLAGNLRENIELGNGPLEAISALYYSTGCTOLPTLARSFICQIMISEAR 215  
 |||||||

QY 171 FOYIEGEMRTIRIRYRRS 188  
 |||||||  
 DB 216 FOYIEGEMRTIRIRYRRS 233  
 |||||||

RESULT 2  
 AGGL\_RICCO STANDARD; PRT; 564 AA.  
 ID AGGL\_RICCO  
 AC P06750;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Agglutinin B chain].  
 OS Ricinus communis (Castor bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae; Eurosid 1; Malpighiales; Euphorbiaceae; Ricinus.  
 NC NCBL\_TaxID=3988;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86059449; PubMed=2999130;  
 RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;  
 RT "The primary sequence of Ricinus communis agglutinin. Comparison with ricin";  
 RL J. Biol. Chem. 260:15682-15686(1985).  
 RN (2)  
 RP SEQUENCE OF 303-564.  
 RC TISSUE=Seed;  
 RA Araki T., Yoshioke Y., Funatsu G.;  
 RT "The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds.";  
 RL Biochim. Biophys. Acta 872:277-285(1986).  
 RN (3)  
 RP SEQUENCE OF 303-337.  
 RX MEDLINE=80178723; PubMed=6768555;  
 RA Lin T.T.-S., Li S.S.-L.;  
 RT "Purification and physicochemical properties of ricins and agglutinins from Ricinus communis";  
 RL Eur. J. Biochem. 105:453-459(1980).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: M12089; AAA33869.1; -  
 DR EMBL: S40368; AAB23584.1; -

DR PIR: A24261; RLCSAG.  
 DR HSSP: P02879; IIR6.  
 DR GlycosylatedDB: P06750;  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PRO0396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS00231; RICIN\_B\_LLECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein; Lectin; Signal.

FT CHAIN 1 24  
 FT SIGNAL 25 290  
 FT PROPEP 291 302  
 FT CHAIN 303 564  
 FT DOMAIN 309 436  
 FT DOMAIN 439 563  
 FT REPEAT 319 361  
 FT REPEAT 362 402  
 FT REPEAT 405 437  
 FT REPEAT 450 485  
 FT REPEAT 489 528  
 FT REPEAT 531 558  
 FT ACT\_SITE 200 200  
 FT DISULFID 282 306  
 FT DISULFID 322 341  
 FT DISULFID 365 382  
 FT DISULFID 453 466  
 FT DISULFID 492 509  
 FT CARBOHYD 34 34  
 FT CARBOHYD 259 259  
 FT CARBOHYD 397 397  
 FT CARBOHYD 437 437  
 FT CARBOHYD 437 437  
 FT CONFLICT 331 331  
 FT CONFLICT 362 362  
 FT CONFLICT 374 374  
 FT CONFLICT 404 404  
 FT CONFLICT 552 552  
 SQ SEQUENCE 564 AA; 62851 MW; D455F2A72F609759 CRC64;

Query Match 89.2%; Score 861.5; DB 1; Length 564;  
 Best Local Similarity 86.9%; Pred. No. 4e-72;  
 Matches 172; Conservative 7; Mismatches 8; Indels 11; Gaps 2;

QY 1 IFFKQYPIINFTAGATVOSTNFIKAVRGRLT-----VLPNRVGLPINORITLV 50  
 |||||||  
 DB 25 IFFKQYPIINFTAGATVOSTNFIKAVRGRLTGTGADVREHETPLVLPNRVGLPINORITLV 84  
 |||||||

QY 51 ELSNHAELSVTLALDVTNAYVYGRAGNSAFFHPDNOEDAETHTLFTDVONRTYPAFG 110  
 |||||||  
 DB 85 ELSNHAELSVTLALDVTNAYVYGRAGNSAFFHPDNOEDAETHTLFTDVONRTYPAFG 144  
 |||||||

QY 111 GNYDRLEQLAGNLRENIELGNGPLEAISALYYSTGCTOLPTLARSFICQIMISEAR 170  
 |||||||  
 DB 145 GNYDRLEQLAGNLRENIELGNGPLEAISALYYSTGCTOLPTLARSFICQIMISEAR 203  
 |||||||

QY 171 FOYIEGEMRTIRIRYRRS 188  
 |||||||  
 DB 204 FOYIEGEMRTIRIRYRRS 221  
 |||||||

RESULT 3  
 RIPT\_TRIKI STANDARD; PRT; 289 AA.  
 ID RIPT\_TRIKI  
 AC P09989;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ribosome-inactivating protein alpha-trichosanthin precursor (rRNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS).  
 OS Trichosanthes kiliowii (Mongolian snake-gourd).

CC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eustosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 CC NCBI\_TaxId=3677;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=Maximowicz;  
 CC MEDLINE=91153657; PubMed=1999291;  
 CC Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;  
 CC "Cloning of trichosanthin cDNA and its expression in *Escherichia*  
 CC coli";  
 CC Gene 97:267-272(1991).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=Maximowicz; TISSUE=Leaf;  
 CC MEDLINE=90256790; PubMed=2341400;  
 CC Chow T., Feldman R.A., Lovett M., Platak M.;  
 CC "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a  
 CC type I ribosome-inactivating protein";  
 CC J. Biol. Chem. 265:8670-8674(1990).  
 CC [3]  
 CC SEQUENCE OF 24-270.  
 CC STRAIN=Maximowicz; TISSUE=Tuberous root;  
 CC MEDLINE=90256789; PubMed=2341399;  
 CC Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,  
 CC Wu P., Hwang K., Platak M.;  
 CC "Primary amino acid sequence of alpha-trichosanthin and molecular  
 CC models for abrin A-chain and alpha-trichosanthin";  
 CC J. Biol. Chem. 265:8665-8669(1990).  
 CC [4]  
 CC SEQUENCE OF 24-270.  
 CC TISSUE=Tuberous root;  
 CC Wang Y., Qian R.Q., Jin S.W., Zhang L.Q., Xia Z.X.,  
 CC Tian G.Y., Ni C.Z.;  
 CC "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and  
 CC application";  
 CC Pure Appl. Chem. 58:789-798(1986).  
 CC [5]  
 CC X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).  
 CC MEDLINE=94344957; PubMed=8066085;  
 CC Zhou F., Fu Z., Chen M., Lin Y., Pan K.;  
 CC "Structure of trichosanthin at 1.88-A resolution";  
 CC Proteins 19:4-13(1994).  
 CC [6]  
 CC X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
 CC MEDLINE=95344383; PubMed=7619070;  
 CC Huang Q., Liu S., Tang Y., Jin S., Wang Y.;  
 CC "Studies on crystal structures, active-centre geometry and  
 CC depurinating mechanism of two ribosome-inactivating proteins";  
 CC Biochem. J. 309:285-298(1995).  
 CC -I- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS  
 CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT  
 CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.  
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC DR EMBL, M34858; AAA34207.1; -;  
 CC DR EMBL, J05434; AAA34206.1; -;  
 CC DR PIR, J70566; RLTZT.  
 CC DR PDB, 1MRJ; 07-FEB-95.  
 CC DR PDB, 1MRK; 07-FEB-95.  
 CC DR PDB, 1TCS; 10-JUL-95.  
 CC DR PDB, 1J4G; 28-JAN-03.

DR PDB, 1NL1; 21-JAN-03.  
 DR PDB, 1OD2; 24-APR-00.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin; Signal; 3D-structure.  
 FT CHAIN 1 23  
 FT SIGNAL 24 270  
 FT PROPEP 271 289  
 FT ACT\_SITE 183 183  
 FT CONFLICT 57 60  
 FT CONFLICT 82 84  
 FT CONFLICT 87 87  
 FT CONFLICT 92 92  
 FT CONFLICT 133 144  
 FT CONFLICT 143 144  
 FT CONFLICT 196 196  
 FT CONFLICT 215 216  
 FT CONFLICT 231 231  
 FT CONFLICT 234 234  
 FT CONFLICT 246 266  
 FT CONFLICT 247 247  
 FT STRAND 25 28  
 FT TURN 30 31  
 FT HELIX 34 46  
 FT TURN 47 47  
 FT STRAND 50 54  
 FT TURN 55 56  
 FT STRAND 57 60  
 FT HELIX 66 69  
 FT STRAND 70 76  
 FT TURN 78 79  
 FT STRAND 82 88  
 FT TURN 89 92  
 FT STRAND 93 99  
 FT TURN 100 101  
 FT STRAND 102 105  
 FT HELIX 109 114  
 FT TURN 115 117  
 FT TURN 120 121  
 FT STRAND 124 127  
 FT HELIX 134 141  
 FT TURN 142 142  
 FT HELIX 145 147  
 FT STRAND 150 150  
 FT HELIX 152 163  
 FT TURN 164 165  
 FT HELIX 167 180  
 FT TURN 181 181  
 FT HELIX 182 186  
 FT STRAND 187 187  
 FT HELIX 188 195  
 FT TURN 196 196  
 FT STRAND 202 202  
 FT TURN 206 226  
 FT HELIX 227 230  
 FT STRAND 231 239  
 FT TURN 241 242  
 FT STRAND 245 250  
 FT TURN 251 252  
 FT HELIX 254 258  
 FT TURN 259 259  
 FT STRAND 260 260  
 FT STRAND 263 263  
 FT TURN 266 268  
 SQ SEQUENCE 289 AA; 31676 MD; 5CE09BB630575BB9 CRC64;  
 Query Match 34.8%; Score 336; DB 1; Length 289;  
 Best Local Similarity 38.9%; Pred. No. 7e-24;  
 Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY 9 INFTAGATVOSTYNTFRVGRGLTLPN-----RVGLPINQRFIVELSNHAEI 58  
 DB 25 VEFRLSGATSSSYGYFISNLR---KALPRERKLYDIPILRLSPQSRALHILFNVADE 81  
 QY 59 SYTLADVTNNAVYVGRAGNSAYFFHPDQEDAE-FAITHLTVDONRYTPAEGCNDRL 117  
 DB 82 TISVADVTNNAVYVGRAGDTSYFF---NEASATEAAKYVFDKARKYTLPLPSGNYERLQ 138  
 QY 118 QIAGNLRENIEELNGPLEEASALYYSTGQTLPLARSPFIICOMISEARFOYIEG 177  
 DB 139 TAAGKIRENIPUGLPLADSAITTLTYFNAN-----SAASALMWLIQSTSEARKYFIEQ 193  
 QY 178 MKTRI 182  
 DB 194 IGRV 198  
 RESULT 4  
 ABRC\_ABRPR STANDARD; PRT; 562 AA.  
 AC P28590;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ABRIN-C precursor [Contains: ABRIN-C A chain (rRNA N-glycosidase)  
 (EC 3.2.2.22); ABRIN-C B chain].  
 OS Abrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 OX NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RX MEDLINE=91266957; PubMed=2050149;  
 RA Wood K.A., Lord J.M., Mawzynczak E.J., Platak M.;  
 RT "Preprobrin: genomic cloning, characterisation and the expression of  
 the A-chain in *Escherichia coli*.";  
 RL Eur. J. Biochem. 198; 723-732(1991).  
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
 B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE  
 BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-  
 INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: X55667; CAA39202.1; .  
 DR PIR: S16022; S16022.  
 DR HSSP: P11140; IABR.  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;

KW Glycoprotein; Lectin; Signal; Pyroglutamate carboxylic acid.  
 FT SIGNAL 1 34  
 FT CHAIN 35 285 ABRIN C A CHAIN (BY SIMILARITY).  
 FT PEPTIDE 286 295 LINKER PEPTIDE (BY SIMILARITY).  
 FT CHAIN 296 562 ABRIN C B CHAIN (BY SIMILARITY).  
 FT DOMAIN 307 434 RICIN B-TYPE LECTIN 1.  
 FT DOMAIN 317 359 RICIN B-TYPE LECTIN 2.  
 FT REPEAT 360 400 1-BETA.  
 FT REPEAT 403 435 1-GAMMA.  
 FT REPEAT 448 483 2-ALPHA.  
 FT REPEAT 487 526 2-BETA.  
 FT REPEAT 529 562 2-GAMMA.  
 FT ACT\_SITE 198 198 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 281 303 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 320 339 BY SIMILARITY.  
 FT DISULFID 363 380 BY SIMILARITY.  
 FT DISULFID 451 464 BY SIMILARITY.  
 FT DISULFID 490 507 BY SIMILARITY.  
 FT MOD\_RES 35 PYROGLUTAMATE CARBOXYLIC ACID  
 FT FT (BY SIMILARITY).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 562 AA; 62817 MW; 1FD0ABC7D7BA6278 CRC64;  
 Query Match 33.9%; Score 327.5; DB 1; Length 562;  
 Best Local Similarity 43.9%; Pred. No. 9,56-23;  
 Matches 82; Conservative 22; Mismatches 68; Indels 15; Gaps 4;  
 QY 5 QYPIINFTAGATVOSTYNTFRVGRGLT-----VLPNRGLPINQRFIVELSNHAEI 56  
 DB 35 ODVVKFTTEGATSSQSYKQFTEALRQRLTGLIDIVLPDPTVEERNRIVTELSENSE 94  
 QY 57 ELSTLADVTNNAVYVGRAGNSAYFFHPDQEDAE-FAITHLTVDONRYTPAEGCNDRL 116  
 DB 95 RESIEVGIDVTNNAVYVGRAGNSQSYFF---RDAPASASTYFPEQTO-RYSLRFPGSYGDL 150  
 QY 117 EQLGNLRENIEELNGPLEEASALYYSTGQTLPLARSPFIICOMISEARFOYIEG 176  
 DB 151 ERMAQHTREELSLGIALTHAIS---FLRSASNDKARFLYIIMASAPARYISN 207  
 QY 177 EMKTRIR 183  
 DB 208 RVGVISIR 214  
 RESULT 5  
 RIPS\_TRIKI STANDARD; PRT; 289 AA.  
 ID RIPS\_TRIKI  
 AC P24478;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein karasurin precursor (rRNA  
 N-glycosidase) (EC 3.2.2.22).  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 OX NCBI\_TaxID=3677;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Root tuber;  
 RX MEDLINE=97356562; PubMed=9212998;  
 RA Mizukami H., Iida K., Kondo T., Ogihara Y.;  
 RT "Cloning and bacterial expression of a gene encoding ribosome-  
 inactivating proteins, karasurin-A and karasurin-C, from *Trichosanthes*  
 kirilowii var. *japonica*.";  
 RL Biol. Pharm. Bull. 20:711-713(1997).  
 RN [2]  
 RP SEQUENCE OF 24-270.  
 RX MEDLINE=92005921; PubMed=1914000;



FT CONFLICT 291 291 D -> N (IN REF. 2).  
 FT CONFLICT 350 351 AE -> PQ (IN REF. 2).  
 FT CONFLICT 378 378 S -> N (IN REF. 2).  
 FT CONFLICT 426 426 L -> M (IN REF. 2).  
 FT CONFLICT 428 428 Y -> D (IN REF. 2).  
 FT CONFLICT 431 431 N -> S (IN REF. 2).  
 FT CONFLICT 484 484 R -> S (IN REF. 2).  
 FT CONFLICT 491 491 H -> Y (IN REF. 2).  
 FT CONFLICT 493 493 H -> G (IN REF. 2).  
 FT CONFLICT 502 502 R -> G (IN REF. 2).  
 FT CONFLICT 509 509 E -> Q (IN REF. 2).  
 FT CONFLICT 513 513 H -> W (IN REF. 2).  
 FT CONFLICT 516 516 H -> T (IN REF. 2).  
 SQ SEQUENCE 527 AA; 59114 MW; 3253AE90CE9494A CRC64;

Query Match 33.4%; Score 323; DB 1; Length 527;  
 Best Local Similarity 43.9%; Pred. No. 2.3e-22;

Matches 83; Conservative 21; Mismatches 69; Indels 16; Gaps 4;

QY 5 QYPIINFTAGATVOSTYNTFIRAVGRRLT-----VLPNRVGLPINQRFILVELSNHA 56  
 DB 1 QDOVIFKFTGATGTSQSYKOFIALRQLTGLIHGIPVLPDPFTLOENRRTYSVELNSD 60  
 QY 57 ELSTALDVNAYVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVQNRRTFAFGNDRL 116  
 DB 61 TESIEAGIDVSNAYVAYRACNRSYEL--RDAPTASRXYLETGTQ-QYSLFNGSYIDL 116  
 QY 117 EOLAGLRNRELNGPLLEASIALYYSTGTOLPLTARSFTICOMISEARROYIEG 176  
 DB 117 ENKARTKPLPLGLQALRHAIISFL--QSGTDQDEIARLTLIVLQMSAEARFYLSY 172  
 QY 177 EMKTRIRYN 185  
 DB 173 RVGVSIKRN 181

RESULT 7  
 RIP2\_BRYDI STANDARD; PRT; 282 AA.

ID RIP2\_BRYDI STANDARD; PRT; 282 AA.  
 AC P98184; 09S8J0;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein bryodin II precursor (rRNA N-glycosidase) (EC 3.2.2.22) (BD2).  
 OS Bryonia dioica (Red bryony).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid 1; Cucurbitales; Cucurbitaceae; Bryonia.  
 ON NCBI\_TaxID=3652;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Siegel C.B., Gawlak S.L., Marguardt H.;  
 RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant Bryonia dioica."  
 RL Patent number US5597569, 28-JAN-1997.  
 RN [2]  
 RP SEQUENCE OF 22-42.  
 RC TISSUE=Root;  
 RX MEDLINE=95151812; PubMed=7849072;  
 RA Siegel C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B., Marguardt H.;  
 RT "Characterization of ribosome-inactivating proteins isolated from Bryonia dioica and their utility as carcinoma-reactive immunokonjugates."  
 RL Bioconjug. Chem. 5:423-429(1994).  
 CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 1 RIP SUBFAMILY.  
 CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/or\\_send\\_an\\_email\\_to\\_license@isb-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch)).

DR EMBL; 134238; -; NOT\_ANNOTATED\_CDS.

DR HSSB; P09989; IMRJ.

DR InterPro; IPR001574; RIP.

DR Pfam; PF00161; RIP.

DR PRINTS; PR00396; SHIGARICIN.

DR PROSITE; PS00275; SHIGA\_RICIN. 1.

KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;

KW Multigene family; Glycoprotein; Signal.

FT SIGNAL 1 21

FT CHAIN 22 282 RIBOSOME-INACTIVATING PROTEIN BRYODIN II.

FT ACT\_SITE 183 183 BY SIMILARITY.

FT CARBOHYD 25 25 N-LINKED (GLCNAC... ) (POTENTIAL).

SQ SEQUENCE 282 AA; 30754 MW; C52BE2FE6A873769C CRC64;

Query Match 33.2%; Score 320.5; DB 1; Length 282;  
 Best Local Similarity 46.0%; Pred. No. 1.8e-22;

Matches 81; Conservative 24; Mismatches 54; Indels 17; Gaps 7;

QY 9 INFTAGATVOSTYNTFIRAVGRRLT-LPNRVGLPINQ-----RPIVELSNHAELSV 60  
 DB 24 INFSLIGATGATYFTFRNLKTLTGTPRYVDIPVLRNAAAGARQLVTLNRNSES 83  
 QY 61 TLALDVNAYVVGVRAGNSAYFFHPDNOEDAEATHTLFTDVQNRRTFAFGNDRL 120  
 DB 84 TVALDVNAYVVGVRAGNTAYFL--ADASTEANNVFAAGI-NHVRPLPGGNDGLETAA 139  
 QY 121 GNL-RENTLGNGLLEASIALYYSTGTOLPLTARSFTICOMISEARROYIEG 175  
 DB 140 GRISRENTLGFSEISAIIGNMFRHP-GTSVP--RAFTVITDTVEAARFKIE 191

RESULT 8

NIGB\_SAMNI STANDARD; PRT; 563 AA.

ID NIGB\_SAMNI STANDARD; PRT; 563 AA.  
 AC P33183; P33184; P93542;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Nigirin b precursor (Agglutinin V) (SNAY) [contains: Nigirin b A chain (rRNA N-glycosidase) (EC 3.2.2.22); Nigirin b B chain].  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 ON NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Van Damme E.J., Barre P., Van Leuven F., Peumans W.J.;  
 RT "Characterization and molecular cloning of Sambucus nigra agglutinin V (nigirin b), a GalNAc-specific type-2 ribosome-inactivating protein from the bark of elderberry (Sambucus nigra)."  
 RL Eur. J. Biochem. 237:505-513(1996).  
 RN [2]  
 RP SEQUENCE OF 26-49 AND 298-321.  
 RC TISSUE=Bark;  
 RX MEDLINE=94003077; PubMed=8400135;  
 RA Githes T., Cliftore L., Ferreres J.M., Rojo M.A., Iglesias R., Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;  
 RT "Isolation and partial characterization of nigirin b, a non-toxic novel type 2 ribosome-inactivating protein from the bark of Sambucus nigra L."  
 RL Plant Mol. Biol. 22:1181-1186(1993).  
 CC -1- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN



CC PROTEIN SYNTHESIS BUT DOES NOT AFFECT PLANT NOR BACTERIAL PROTEIN  
 CC SYNTHESIS. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
 CC B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE  
 CC BINDING OF NIGRIN B TO THE CELL MEMBRANE THAT PRECEDES  
 CC ENDOCYTOSIS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.  
 CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, U41299; AAB39475.1; -.  
 DR PIR, S37382; S37382.  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 DR Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; Signal.  
 FT SIGNAL 1 25  
 FT CHAIN 26 297  
 FT CHAIN 298 563  
 FT DOMAIN 305 431  
 FT DOMAIN 434 559  
 FT REPEAT 316 356  
 FT REPEAT 357 397  
 FT REPEAT 400 432  
 FT REPEAT 445 482  
 FT REPEAT 486 524  
 FT REPEAT 527 554  
 FT ACT\_SITE 188 188  
 FT DISULFID 274 302  
 FT DISULFID 319 338  
 FT DISULFID 360 377  
 FT DISULFID 448 463  
 FT DISULFID 489 506  
 FT CARBOHYD 221 221  
 FT CARBOHYD 368 368  
 FT CARBOHYD 376 376  
 FT CARBOHYD 483 483  
 FT CARBOHYD 537 537  
 FT CONFLICT 39 39  
 SQ SEQUENCE 563 AA; 62300 MW; F250CBE24621BF14 CRC64;  
 Query Match 31.8%; Score 307.5; DB 1; Length 563;  
 Best Local Similarity 38.4%; Pred. No 6,7e-21;  
 Matches 73; Conservative 36; Mismatches 56; Indels 25; Gaps 5;  
 QY 6 YPIINFTTGAATVOSTYNTFRVAVR-----GRITVLNPNVGLPIINORFLIVELSNH 55  
 Db 28 YPSVFMUDGAKSATSVDLFLSNLRKTVATGTYEVNGLPVLRSREVSQVSRFVLVPLTNY 87  
 QY 56 AELSVTLALDVNTANVYVVGRAAGSAYFFHPDNDDEAAT--TLIFDVQNRITFAAGCY 113  
 Db 88 NGNVTTLAVDVNTNLYVVAFGSNNASYFF-----KDATVYOKSNLFVGTQKN-TLSEFGNY 141  
 QY 114 DRLEQLGNLRNENLELNGPLEFAISALYYSTGGQLDPLTASSFICIMISEARFOY 173

Db 142 DNLEFANTRRSTELGSPIDCAITSLYHGD-----SVARSLLVIQVSEARFRY 194  
 QY 174 IEGEMRTAIR 183  
 Db 195 IEQEVRSIQ 204  
 RESULT 9  
 ID RIP1\_BRD1 STANDARD; PRT; 290 AA.  
 AC P33185; Q95819;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein bryodin I precursor (rRNA N-glycosylase)  
 DE (EC 3.2.2.22) (BD1).  
 OS Bryonia dioica (red bryony).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.  
 OX NCBI\_TaxID=3652;  
 [1]  
 RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
 RC TISSUE=Leaf;  
 RA Gawiak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,  
 RA Siegall C.B.;  
 RT "Molecular, biological, and preliminary structural analysis of  
 RT recombinant bryodin I, a ribosome-inactivating protein from the plant  
 RT Bryonia dioica."  
 RL Biochemistry 36:3095-3103(1997).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RA Siegall C.B.;  
 RT "Cloning and expression of a gene encoding bryodin I from Bryonia  
 RT dioica."  
 RL Patent number US5541110, 30-JUL-1996.  
 [3]  
 RP SEQUENCE OF 24-66.  
 RC TISSUE=Seed;  
 RA MEDLINE=89326691; PubMed=2753596;  
 RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Sorla M.,  
 RA Lappi D.;  
 RT "N-terminal sequence of some ribosome-inactivating proteins."  
 RL Int. J. Pept. Protein Res. 33:263-267(1989).  
 [4]  
 RP SEQUENCE OF 24-43.  
 RC TISSUE=Root;  
 RA MEDLINE=95151812; PubMed=7849072;  
 RA Siegall C.B., Gawiak S.L., Chace D., Wolff E.A., Mixan B.,  
 RA Marguardt H.;  
 RT "Characterization of ribosome-inactivating proteins isolated from  
 RT Bryonia dioica and their utility as carcinoma-reactive  
 RT immunoconjugates."  
 RL Bioconj. Chem. 5:423-429(1994).  
 CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
 CC PROTEIN SYNTHESIS IN ANIMAL CELLS.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO  
 CC PRODUCE A SHORTER PROTEIN.  
 CC -1- BIOTECHNOLOGY: Especially useful as immunotoxin for  
 CC pharmacological applications as it has low toxicity in rats and  
 CC mice but is potent once inside target cells.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY,  
 CC TYPE 1 RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>



```

or send an email to license@sb-sib.ch).
-----
CC EMBL, I24020.  NOT_ANNOTATED_CDS.
DR PIR, S16491; S16491.
DR PDB; 1BRX; 04-MAR-98.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN. 1.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
  3D-structure; Multigene family; Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 1 270 RIBOSOME-INACTIVATING PROTEIN BRYODIN I.
FT PROPE 271 290 MISSING IN MATURE PROTEIN.
FT ACT_SITE 183 183 BY SIMILARITY.
FT ACT_SITE 212 212
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MUTAGEN 212 212 E->K; REDUCES ACTIVITY 10-FOLD.
FT CONFLICT 61 65 RSSIS -> LRHXI (IN REF. 3).
FT STRAND 25 28
FT TURN 30 31
FT TURN 34 46
FT HELIX 47 47
FT TURN 50 54
FT STRAND 55 56
FT TURN 57 60
FT HELIX 66 69
FT STRAND 70 76
FT TURN 78 79
FT STRAND 82 88
FT TURN 89 92
FT STRAND 93 99
FT TURN 100 101
FT STRAND 102 105
FT HELIX 109 114
FT TURN 115 117
FT TURN 120 121
FT STRAND 124 127
FT HELIX 134 141
FT TURN 142 142
FT HELIX 145 147
FT STRAND 150 150
FT HELIX 152 163
FT TURN 164 165
FT HELIX 167 186
FT STRAND 187 187
FT HELIX 188 196
FT STRAND 202 202
FT HELIX 206 213
FT TURN 214 214
FT HELIX 215 225
FT TURN 226 230
FT STRAND 231 239
FT TURN 241 242
FT STRAND 245 250
FT TURN 251 252
FT HELIX 254 257
FT TURN 258 259
FT STRAND 260 260
FT STRAND 263 263
FT HELIX 266 268
SQ SEQUENCE 290 AA; 31788 MW; E966CD9C031A42DB CRC64;
Query Match 31.6%; Score 305; DB 1; Length 290;
Best Local Similarity 35.7%; Pred. No. 5.2e-21;
Matches 65; Conservative 45; Mismatches 56; Indels 16; Gaps 4;
OY 9 INFTGATVOSTNIRAVRGRLTVLPNRVGLPI-----NORFIIVELSNHAELSVT 61
DB 25 VSRRLSGATTTSVGVFKRLRELPERKRVYINPLIRSISSGRTLLHLTLNVADETIS 84
OY 62 LALDVTNAVYVGVRAGNSAYFFHPDQEDA-EAITHLFTDVONRYTFAFGNVDRLLEOLA 120

```

```

Db 85 VAADVTVNVMYIMGLADGVSYFF---NEASATEAAKVFVEDAKKRYTLPSGMYERLQTPA 141
OY 121 GNARENIEGNGPFLERNAISALYYISGCGQLTLARSFIFICIONISEARFOYIEGKRT 180
Db 142 GKIRENIPGLPALDSATITLLYYTAS-----SASALLVLIQSTAESARKYFIEDQIGK 196
OY 181 RI 182
Db 197 RV 198

RESULT 10
ABRA_ABRPR STANDARD; PRT; 528 AA.
ID ABRA_ABRPR
AC P11140; P28589;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Abirin-a precursor [Contains: Abirin-a A chain (rRNA N-glycosidase)
  (EC 3.2.2.22); Abirin-a B chain].
OS Abrus precatorius (Indian Licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_Taxid=3816;
RN [1]
RP MEDLINE=93132798; PubMed=8421313;
RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
RT "Primary structure of three distinct Isoabirins determined by cDNA
  sequencing. Conservation and significance."
RL J. Mol. Biol. 229:263-267(1993).
RN [2]
RP SEQUENCE OF 1-251.
RC TISSUE=Seed;
RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
RT "The complete amino acid sequence of the A-chain of abirin-a, a toxic
  protein from the seeds of Abrus precatorius."
RL Agric. Biol. Chem. 52:1095-1097(1988).
RN [3]
RP SEQUENCE OF 1-251 FROM N.A.
RC TISSUE=Leaf;
RA Evensen G., Mathiesen A., Sundan A.;
RT "Direct molecular cloning and expression of two distinct abirin
  A-chains."
RL J. Biol. Chem. 266:6848-6852(1991).
RN [4]
RP SEQUENCE OF 262-528.
RA Chen Y.-L., Chow L.-P., Tsungta A., Lin J.-Y.;
RT "The complete primary structure of abirin-a B chain."
RL FEBS Lett. 309:115-118(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).
RA MEDLINE=95333188; PubMed=7608980;
RA Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
RT "Crystal structure of abirin-a at 2.14 A."
RL J. Mol. Biol. 250:354-367(1995).
CC -I- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
  SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
  SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC ABRIN-A IS MORE TOXIC THAN RICIN.
CC -I- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
  FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
  PRECEDES ENDOCYTOSIS.
CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
  specific adenosine on the 28S rRNA.
CC -I- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -I- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
  CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -I- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-

```

INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
-1- SIMILARITY: Contains 2 ricin B-type lectin domains.  
-----  
CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: M98344; AAA32624.1; ALT\_INIT.  
DR EMBL: X54872; ; NOT\_ANNOTATED\_CDS.  
DR PIR: S32429; TLISA.  
DR PDB: 1ABR; 07-FEB-95.  
DR InterPro: IPR000772; Ricin\_B\_lectin.  
DR InterPro: IPR001574; RIP.  
DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR SMART: SM00458; RICIN; 2.  
DR PROSITE: PS50231; RICIN\_B\_LLECTIN; 2.  
DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
KW Glycoprotein; Lectin; 3D-structure; Pyrrolidone carboxylic acid.  
FT CHAIN 1 251 ABRIN-A A CHAIN.  
FT PEPTIDE 252 261 LINKER PEPTIDE.  
FT CHAIN 262 528 ABRIN-A B CHAIN.  
FT DOMAIN 273 400 RICIN B-TYPE LECTIN 1.  
FT DOMAIN 403 527 RICIN B-TYPE LECTIN 2.  
FT REPEAT 283 325 1-ALPHA.  
FT REPEAT 326 366 1-BETA.  
FT REPEAT 369 401 1-GAMMA.  
FT REPEAT 414 449 2-ALPHA.  
FT REPEAT 453 492 2-BETA.  
FT REPEAT 495 528 2-GAMMA.  
FT ACT\_SITE 164 164 BY SIMILARITY.  
FT DISULFID 247 269 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 286 305 BY SIMILARITY.  
FT DISULFID 329 346 BY SIMILARITY.  
FT DISULFID 417 430 BY SIMILARITY.  
FT DISULFID 456 473 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 1 1 BY SIMILARITY.  
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .).  
FT CONFLICT 202 202 MISSING (IN REF. 2).  
FT CONFLICT 298 298 N->Y (IN REF. 4).  
FT CONFLICT 427 427 M->L (IN REF. 4).  
FT CONFLICT 467 467 T->P (IN REF. 4).  
FT CONFLICT 483 483 V->L (IN REF. 4).  
FT STRAND 5 8  
FT TURN 10 11  
FT HELIX 14 28  
FT STRAND 32 33  
FT TURN 34 35  
FT STRAND 36 38  
FT TURN 42 43  
FT HELIX 47 49  
FT STRAND 51 57  
FT STRAND 63 69  
FT TURN 72 79  
FT STRAND 75 79  
FT STRAND 83 86  
FT TURN 88 89  
FT TURN 92 93  
FT TURN 94 97  
FT HELIX 100 101  
FT STRAND 103 106  
FT TURN 113 114  
FT HELIX 115 119  
FT TURN 124 126  
FT STRAND 129 129  
FT HELIX 131 142

FT TURN 143 144  
FT HELIX 148 167  
FT STRAND 168 168  
FT HELIX 169 180  
FT TURN 181 182  
FT STRAND 185 185  
FT HELIX 189 196  
FT TURN 197 197  
FT HELIX 198 207  
FT STRAND 207 220  
FT TURN 212 222  
FT TURN 222 223  
FT STRAND 226 231  
FT TURN 232 233  
FT HELIX 235 239  
FT STRAND 240 240  
FT STRAND 243 243  
FT STRAND 248 248  
FT STRAND 268 268  
FT STRAND 276 277  
FT STRAND 279 280  
FT HELIX 282 284  
FT STRAND 286 289  
FT HELIX 290 292  
FT TURN 292 297  
FT STRAND 299 303  
FT STRAND 303 313  
FT HELIX 311 317  
FT STRAND 315 317  
FT TURN 319 320  
FT STRAND 322 325  
FT TURN 326 327  
FT STRAND 328 332  
FT TURN 337 338  
FT STRAND 340 344  
FT TURN 346 348  
FT HELIX 351 353  
FT STRAND 355 355  
FT STRAND 357 358  
FT TURN 360 361  
FT STRAND 364 366  
FT TURN 367 370  
FT STRAND 371 374  
FT TURN 380 381  
FT STRAND 383 383  
FT STRAND 385 387  
FT HELIX 393 395  
FT STRAND 398 399  
FT STRAND 406 408  
FT STRAND 410 411  
FT HELIX 413 415  
FT STRAND 417 421  
FT TURN 422 423  
FT STRAND 424 428  
FT TURN 432 433  
FT HELIX 435 437  
FT STRAND 439 441  
FT TURN 443 444  
FT STRAND 447 449  
FT TURN 450 451  
FT STRAND 452 459  
FT TURN 464 465  
FT STRAND 467 472  
FT TURN 474 475  
FT HELIX 478 480  
FT STRAND 483 484  
FT TURN 486 487  
FT STRAND 490 492  
FT TURN 493 496  
FT STRAND 497 501

Query Match 31.4%; Score 303.5; DB 1; Length 528;  
Best Local Similarity 41.1%; Pred. No. 1.5e-20;  
Matches 76; Conservative 26; Mismatches 64; Indels 19; Gaps 5;





DR PDB; 1AH4; 22-JUN-94.  
 DR PDB; 1AHB; 22-JUN-94.  
 DR PDB; 1AHC; 22-JUN-94.  
 DR PDB; 1MOM; 31-MAY-94.  
 DR PDB; 1MRG; 07-FEB-95.  
 DR PDB; 1MRH; 07-FEB-95.  
 DR PDB; 1MRI; 07-FEB-95.  
 DR PDB; 1F9Q; 21-NOV-01.  
 DR GlycoSuiteDB; P16094; .  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;  
 KW Glycoprotein; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 269  
 FT PROPP 270 286  
 FT ACT\_SITE 183 183  
 FT CARBOHYD 250 250  
 FT STRAND 25 28  
 FT HELIX 34 47  
 FT STRAND 50 54  
 FT TURN 55 56  
 FT STRAND 57 60  
 FT HELIX 66 69  
 FT STRAND 70 76  
 FT TURN 78 79  
 FT STRAND 82 88  
 FT TURN 85 92  
 FT STRAND 93 99  
 FT TURN 100 101  
 FT STRAND 102 105  
 FT HELIX 109 114  
 FT TURN 115 117  
 FT TURN 120 121  
 FT STRAND 124 127  
 FT HELIX 134 141  
 FT HELIX 145 147  
 FT STRAND 150 150  
 FT STRAND 152 162  
 FT TURN 163 163  
 FT HELIX 167 168  
 FT STRAND 187 187  
 FT HELIX 188 195  
 FT TURN 196 197  
 FT STRAND 202 202  
 FT HELIX 206 225  
 FT TURN 226 230  
 FT STRAND 231 238  
 FT TURN 240 241  
 FT STRAND 246 250  
 FT TURN 251 252  
 FT HELIX 254 258  
 FT TURN 259 259  
 FT STRAND 260 260  
 FT STRAND 263 263  
 FT HELIX 266 266  
 SQ SEQUENCE 286 AA; 31532 MW; E1B013ABEC216CF CRC64;

Query Match 28.3%; Score 273; DB 1; Length 286;  
 Best Local Similarity 34.8%; Pred. No. 4.6e-18;  
 Matches 63; Conservative 40; Mismatches 62; Indels 16; Gaps 4;

OY 9 INFTAGATVQSYNFIKRAVGRGLTVLPNRVGLPI-----NORFIIVELSNHAEISVT 61  
 DB 25 VEFRLSGADPRSGYKIKRLNALPPREKVVYIPIILLSVSAGRYLMLHLEFNYSCKIT 84  
 OY 62 LADVVTNAVYVGYRAGNSAVFFHPDQEDAE-ATHTLEFDVQNRVTFAGGNYDRLEQLA 120  
 DB 85 VADVVTNAVYVGYRAGNSAVFFHPDQEDAE-ATHTLEFDVQNRVTFAGGNYDRLEQLA 141

OY 121 GNLRNIEIENGNGPLEEASIALYVSTGCTQLPTLARSPICIQIMSEAPRQYIEGMRT 180  
 DB 142 CKPREKIRIGUPALDSATSTLLHDS-----TAAAGALLVLIQTAAERKRYIEQIQE 196  
 OY 181 R 181  
 DB 197 R 197

RESULT 15  
 RIP2\_MOMBA STANDARD; PRT; 286 AA.  
 AC P29339;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ribosome-inactivating protein momordin II precursor (rRNA N-glycosidase) (BC 3.2.2.22).  
 OS Momordica balsamina (Bitter melon) (Balsam pear).  
 OC Momordica; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Eukaryota; Viridiplantae; eudicotyledons; core eudicot; Rosidae;  
 OC Eucosida; Cucurbitales; Cucurbitaceae; Momordica.  
 OX NCBI\_TaxID=3672;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seed;  
 RX MEDLINE=93027170; PubMed=1408771;  
 RA Ortigao M., Better M.;  
 RT "Momordin II, a ribosome inactivating protein from Momordica  
 RL Nucleic Acids Res. 20:4662-4662(1992).  
 CC -! CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -! SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; Z12175; CAA78166.1; .  
 DR PIR; S25560; S25560.  
 DR PDB; 1CF5; 07-JUN-99.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 286  
 FT ACT\_SITE 181 181  
 FT SEQUENCE 286 AA; 32031 MW; 3B89EF1AE6B25986 CRC64;

Query Match 28.2%; Score 272; DB 1; Length 286;  
 Best Local Similarity 34.2%; Pred. No. 5.7e-18;  
 Matches 63; Conservative 39; Mismatches 66; Indels 16; Gaps 4;

OY 9 INFTAGATVQSYNFIKRAVGRGLTVLPNRVGLPI-----NORFIIVELSNHAEISVT 61  
 DB 25 VNFPLSTAKATKTYKFLIEDPRATLPFSHKYDIPLXSTSDSRFLDLLDTSYAETIS 84  
 OY 62 LADVVTNAVYVGYRAGNSAVFFHPDQEDAEATHTLEFDVQNRVTFAGGNYDRLEQLA 121  
 DB 85 VADVVTNAVYVGYRAGNSAVFFHPDQEDAEATHTLEFDVQNRVTFAGGNYDRLEQLA 140  
 OY 122 NLRENIEIENGNGPLEEASIALYVSTGCTQLPTLARSPICIQIMSEAPRQYIEGMRT 181

Db 141 KIRENIDLGIPALSAITTLFFYNA-----QSAPSAVLVLIQTAEARFKYIERHVAKY 195

QY 182 IRYN 185  
: |

Db 196 VATN 199

Search completed: September 16, 2003, 11:46:04  
Job time : 7.59071 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:40:08 : Search time 29.9877 Seconds

(without alignments)  
1617.791 Million cell updates/sec

Title: US-10-083-336a-4

Sequence: 1 IFFKQYPIINFITACATVQS.....ARFOYEGEMRTIRIKYRNS 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SPTRMBL\_23:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organellar:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*  
16: sp\_virus:\*  
17: sp\_bacteriapi:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	951	98.4	541	10 Q41174	Q41174 ricinus com
2	374.5	38.8	580	10 Q94BW3	Q94BW3 cinnamomum
3	372.5	38.6	580	10 Q94BW4	Q94BW4 cinnamomum
4	370.5	38.4	581	10 Q94BW5	Q94BW5 cinnamomum
5	368.5	38.1	549	10 Q9FV22	Q9FV22 cinnamomum
6	334	34.6	289	10 Q94KE4	Q94KE4 trichosan
7	332	34.4	289	10 Q41216	Q41216 trichosan
8	329.5	34.1	563	10 Q04367	Q04367 sambucus ni
9	326	33.7	247	10 Q91RE3	Q91RE3 trichosan
10	323.5	33.5	564	10 Q9AVR2	Q9AVR2 sambucus eb
11	318.5	33.0	528	10 Q06076	Q06076 abrus preca
12	317	32.8	270	10 Q8LPU7	Q8LPU7 trichosan
13	310.5	32.1	252	10 Q38760	Q38760 abrus preca
14	307.5	31.8	563	10 Q945S2	Q945S2 sambucus ni
15	307.5	31.8	563	10 Q8GT32	Q8GT32 sambucus ni
16	305	31.6	270	10 Q41611	Q41611 trichosan

17	303.5	31.4	252	10 Q38761	Q38761 abrus preca
18	302	31.3	565	10 Q04071	Q04071 sambucus ni
19	300.5	31.1	278	10 Q00980	Q00980 luffa cylin
20	298.5	30.9	547	10 Q9M6E9	Q9M6E9 abrus preca
21	297.5	30.8	251	10 Q96236	Q96236 abrus preca
22	296.5	30.7	251	10 Q96237	Q96237 abrus preca
23	293.5	30.4	566	10 Q04072	Q04072 sambucus ni
24	290.5	30.1	251	10 Q96235	Q96235 abrus preca
25	273	28.3	264	10 Q9F8H2	Q9F8H2 momordica c
26	273	28.3	570	10 Q41358	Q41358 sambucus ni
27	271	28.1	592	10 Q8M2E7	Q8M2E7 iris holian
28	270	28.0	249	10 Q8LK05	Q8LK05 viscum albu
29	270	28.0	286	10 Q9FUV7	Q9FUV7 momordica c
30	267	27.6	570	10 Q22415	Q22415 sambucus ni
31	266.5	27.6	604	10 Q9M654	Q9M654 polygonatum
32	265	27.4	412	10 Q41257	Q41257 momordica c
33	263	27.2	565	10 Q8M243	Q8M243 viscum albu
34	260	26.9	254	10 Q8LK06	Q8LK06 viscum albu
35	260	26.9	573	10 Q8M2E8	Q8M2E8 iris holian
36	259	26.8	251	10 Q8LK04	Q8LK04 viscum albu
37	259	26.8	293	10 Q8S452	Q8S452 jatropa cu
38	259	26.8	569	10 P93543	P93543 sambucus ni
39	253	26.2	531	10 Q8RXH6	Q8RXH6 viscum albu
40	251	26.0	249	10 Q8RXH7	Q8RXH7 viscum albu
41	247	25.6	258	10 Q9S9E4	Q9S9E4 gellonium mu
42	247	25.6	293	10 Q8VYU0	Q8VYU0 jatropa cu
43	246.5	25.5	203	10 Q8RY69	Q8RY69 gnostemma
44	246.5	25.5	275	10 Q8H1Y4	Q8H1Y4 gnostemma
45	245.5	25.4	603	10 Q9M653	Q9M653 polygonatum

## ALIGNMENTS

RESULT 1  
Q41174 PRELIMINARY: PRT; 541 AA.

AC 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Protein A chain (EC 3.2.2.22) (rRNA N-glycosidase)  
DE (Fragment).  
OS Ricinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3968;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92338377; PubMed=1633311;  
RA Roberts L.M., Tregear J.W., Lord J.M.;  
RT "Molecular cloning of ricin";  
RL Targeted Diagn. Ther. 7:81-97(1992).  
CC -I- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RNA.  
CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
EMBL: S40366; AAB22582.1; -.  
DR HSSP: P02879; IBR6.  
DR InterPro: IPR000772; Ricin\_B\_lectin.  
DR InterPro: IPR001574; RIP.  
DR InterPro: IPR001400; Somatotropin.  
DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
DR Pfam: PF00161; RIP; 1.  
DR PRINTS: PR00396; SHIGARICIN.  
DR SMART: SM00458; RICIN; 2.  
DR PROSITE: PS00231; RICIN\_B\_LECTIN; 2.  
DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
DR PROSITE: PS00338; SOMATOTROPIN\_2; 1.  
KW Hydrolyase; Toxin.  
FT NON\_TER 1  
SQ SEQUENCE 541 AA; 60281 MW; 287B2CDEF1F2E9D9 CRC64;

Query Match 98.4%; Score 951; DB 10; Length 541;  
 Best Local Similarity 94.9%; Pred. No. 2,7e-82;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 IFPKVPIINPTAGATVOSTNIRAVRGRLT-----VLPNRVGLPINORITLY 50  
 DB 1 IFPKVPIINPTAGATVOSTNIRAVRGRLTGGADVDRHDIPVLPNRVGLPINORITLY 60

QY 51 ELSNHAELSVTLADVTNAYVYGRAGNSAYFFHPDNOEDAEATITLFTDVONRYTFAG 110  
 DB 61 ELSNHAELSVTLADVTNAYVYGRAGNSAYFFHPDNOEDAEATITLFTDVONRYTFAG 120

QY 111 GNDRLPOLAGNIRENIELGNGPLEAISALYYSTGTQLPTLARSFICIQMISEAR 170  
 DB 121 GNDRLPOLAGNIRENIELGNGPLEAISALYYSTGTQLPTLARSFICIQMISEAR 180

QY 171 FOYIEGEMRTIRIRNRS 188  
 DB 181 FOYIEGEMRTIRIRNRS 198

RESULT 2  
 Q94BW3 PRELIMINARY; PRT; 580 AA.  
 AC 094BW3:  
 DT 01-DEC-2001 (TRENBLREL. 19, Created)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomlin III precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang O., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomlin proteins and study of their expression  
 RT patterns.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADEOSINE ON THE 28S RRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AY039803; AAK82460.1;  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; Ricin.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 KW Hydrolase; Signal; Toxin.  
 FT SIGNAL 1  
 FT CHAIN 33 580  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMLIN III  
 SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7FB558 CRC64;

Query Match 38.8%; Score 374.5; DB 10; Length 580;  
 Best Local Similarity 48.1%; Pred. No. 3,7e-27;  
 Matches 90; Conservative 26; Mismatches 56; Indels 15; Gaps 5;

QY 6 YPIINFTAGATVOSTNIRAVRGRLT-----VLPNRVGLPINORITLYELSN-H 55  
 DB 33 YQVYFTTKNATKTSYTFIEALRAQLASGEHPGIPVNRKSTVPDSKRTLYELSNWA 92

QY 56 AELSVTLADVTNAYVYGRAGNSAYFFHPDNOEDAEATITLFTDVONRYTFAGGNDYR 115  
 DB 93 ADSPVLAADVNTNAYVYGRAGNSAYFFHPDNOEDAEATITLFTDVONRYTFAGGNDYR 149

QY 116 LEQALAGNIRENIELGNGPLEAISALYYSTGTQLPTLARSFICIQMISEARFOYIE 175  
 DB 150 LERVAGEERREILLGMDPLENAISALMISNL--NQORALARSLIVYQWVAEAVRFRFIE 207

QY 176 GEMRTIR 182  
 DB 208 YRVRESI 214

RESULT 3  
 Q94BW4 PRELIMINARY; PRT; 580 AA.  
 AC 094BW4:  
 DT 01-DEC-2001 (TRENBLREL. 19, Created)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomlin II precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang O., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomlin proteins and study of their expression  
 RT patterns.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADEOSINE ON THE 28S RRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AY039802; AAK82459.1;  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; Ricin.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 KW Hydrolase; Signal; Toxin.  
 FT SIGNAL 1  
 FT CHAIN 33 580  
 FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN  
 FT CINNAMOMLIN II.  
 SQ SEQUENCE 580 AA; 64265 MW; 37E4A289ECCECBFC CRC64;

Query Match 38.6%; Score 372.5; DB 10; Length 580;  
 Best Local Similarity 47.6%; Pred. No. 5,7e-27;  
 Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY 6 YPIINFTAGATVOSTNIRAVRGRLT-----VLPNRVGLPINORITLYELSN-H 55  
 DB 33 YQVYFTTKNATKTSYTFIEALRAQLASGEHPGIPVNRKSTVPDSKRTLYELSNWA 92

QY 56 AELSVTLADVTNAYVYGRAGNSAYFFHPDNOEDAEATITLFTDVONRYTFAGGNDYR 115  
 DB 93 ADSPVLAADVNTNAYVYGRAGNSAYFFHPDNOEDAEATITLFTDVONRYTFAGGNDYR 149

QY 116 LEQALAGNIRENIELGNGPLEAISALYYSTGTQLPTLARSFICIQMISEARFOYIE 175  
 DB 150 LERVAGEERREILLGMDPLENAISALMISNL--WTSNINQORALARSLIVYQWVAEAVRFRFIE 207

QY 176 GEMRTIR 182  
 DB 208 YRVRESI 214

RESULT 4  
 Q94BW5 PRELIMINARY; PRT; 581 AA.  
 AC 094BW5:  
 DT 01-DEC-2001 (TRENBLREL. 19, Created)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomlin I precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).





```

Db      82 TISVAIDVTNYINGRAGDTSYF---NEASATEAAKRYFKDSMRKTLTPYSGNVERLQ 138
QY      118 QLAGNLRNIELGNPLEAISALYXYSTGTOPLTARSTIIICOMISEARQYIEGE 177
Db      139 TAAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEARAKFIEQ 193
QY      178 MRRRI 182
Db      194 IGRKV 198

RESULT 7
Q41216 PRELIMINARY: PRT: 289 AA.
AC Q41216;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Trichosanthin (EC 3.2.2.22) (rRNA N-glycosidase).
GN TRICHOSANTHIN, TCS.
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94271613; PubMed=8003348;
RA Zheng H., Wang B., Shaw P., Yeung H.;
RT "Cloning and DNA sequencing of the gene encoding trichosanthin.";
RL I Chuan Hsueh Pao 21:42-51(1994).
CC -I- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL: S70176; AAC31048.1;
DR HSSP: P09989; IMRJ
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN.1.
KW Hydrolase; Toxin.
SQ SEQUENCE 289 AA; 31650 MW; 286AC1AD48BCA175 CRC64;

Query Match 34.4%; Score 332; DB 10; Length 289;
Best Local Similarity 38.9%; Pred. No. 1.6e-23;
Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY 9 INFTTAGATVQSYNFIKRAVRGRITVLPN-----RVGLPINRFTLVELSNHAE 58
Db 25 VSEPLSGATSSSGVFISSNR--KALPNERKLYDIPLRSSLPGSORVALVHLTNVAD 81
QY 59 SVTLALDVTNAYVVGVRAGNSAVFFHPDNOEDA-EATTHLFTVQNKYTFAGGNDRL 117
Db 82 TISVAIDVTNYINGRAGDTSYF---NEASATEAAKRYFKDSMRKTLTPYSGNVERLQ 138
QY 118 QLAGNLRNIELGNPLEAISALYXYSTGTOPLTARSTIIICOMISEARQYIEGE 177
Db 139 TAAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMWLIQSTSEARAKFIEQ 193
QY 178 MRRRI 182
Db 194 IGRKV 198

RESULT 8
Q04367 PRELIMINARY: PRT: 563 AA.
AC Q04367;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Ribosome inactivating protein precursor (EC 3.2.2.22) (rRNA
DE N-glycosidase).

```

```

OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.
OX NCBI_TaxID=4202;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98112023; PubMed=9450339;
RA Van Damme E.J., Roy S., Barre A., Rouge P., Van Leuven F.,
RA Peumans W.D.;
RT "The major elderberry (Sambucus nigra) fruit protein is a lectin
RT derived from a truncated type 2 ribosome-inactivating protein.";
RL Plant J. 12:1251-1260(1997).
CC -I- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL: U76524; AAC15886.1;
DR HSSP: P02879; ZAAI.
DR InterPro: IPR000772; Ricin_B_lectin.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN.2.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN.1.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 25
FT CHAIN 26 297 POTENTIAL.
FT CHAIN 298 563 RIBOSOME INACTIVATING PROTEIN, A CHAIN.
SQ SEQUENCE 563 AA; 62336 MW; 3ED2B608E796205 CRC64;

Query Match 34.1%; Score 329.5; DB 10; Length 563;
Best Local Similarity 40.5%; Pred. No. 7.1e-23;
Matches 77; Conservative 37; Mismatches 51; Indels 25; Gaps 5;

QY 6 YPIINFTTAGATVQSYNFIKRAVR-----GRITVLPNRVGLPINRFTLVELSNH 55
Db 28 YPSVSEPLSGATSSSGVFISSNR--KALPNERKLYDIPLRSSLPGSORVALVHLTNVAD 81
QY 56 AELSVTLALDVTNAYVVGVRAGNSAVFFHPDNOEDA-EI--THLFTVQNKYTFAGGND 113
Db 88 NGNTVTLADVTVNLVVAASANANSYF---KDTQLQKSNLFLVGR-OHTLPFGCN 141
QY 114 DRLEQLAGNLRNIELGNPLEAISALYXYSTGTOPLTARSTIIICOMISEARQY 173
Db 142 DNLFTAAAGTRRSSTELGSPDLGALTSIYDE-----SVARSLIVYIOMVSEARARY 194
QY 174 IEGEMTRIR 183
Db 195 IEQEVRSIQ 204

RESULT 9
Q9LRE3 PRELIMINARY: PRT: 247 AA.
AC Q9LRE3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Trichobakin (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).
GN TBK.
OS Trichosanthes sp. Bac Kan 8-98.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=118182;
RN [1]
RP SEQUENCE FROM N.A.
RA Phan Van Chi, Nguyen Thuy Ha, Le Tran Binh;
RT "Genomic DNA clone for mature typ-1 ribosome-inactivating protein from
RT Trichosanthes sp. sample 01 Bac Kan 8-98 (Hanoi).";
SQ Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

```

CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: AB039324; BAA92530.1; --  
 DR HSSP: P09989; IMRT.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KM Hydroxylase; Toxin.  
 FT NON\_TER 1 1  
 FT NON\_TER 247 247  
 SQ SEQUENCE 247 AA; 27199 MW; 89811AC32892P03F CRC64;  
 Query Match 33.7%; Score 326; DB 10; Length 247;  
 Best Local Similarity 39.6%; Pred. No. 4.9e-23;  
 Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;  
 QY 9 INFTTAGATVOSTYNTFIRAVR-----GLTVLP-NRVGLPINRFLIVELSNHLSVT 61  
 DB 2 VSFRLSGATSSSYGVFINLKRALPYERKLYDIPLRFSTLPQSQRALHLTNVADETIS 61  
 QY 62 LALDYTNNAVYVGRAGNSAYFFHPNODDA-EATHLFTDYQNRRTFAFGNVDLEOLA 120  
 DB 62 VALDYTNNAVYVGRAGNSAYFFHPNODDA-EATHLFTDYQNRRTFAFGNVDLEOLA 118  
 QY 121 GNLRNIELGNPLLEAISAALYYSTGGTQPLTLARSPFLICIONISEARFOYIEGEMRT 180  
 DB 119 KXIRENIPGLPALDSALTITLTFYVYAN-----SAASALMWLIQSTSEARAKFIDQIGK 173  
 QY 181 RT 182  
 DB 174 RV 175  
 RESULT 10  
 Q9AVR2 PRELIMINARY; PRT; 564 AA.  
 AC Q9AVR2;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Ribosome-inactivating protein precursor (EC 3.2.2.22) (rRNA  
 DE N-glycosidase).  
 GN EBU1.  
 OS Sambucus ebulus.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=28503;  
 RN NCB1\_TaxID=28503;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=leaf;  
 RA Clrbes T., Iglesias R., Perez Y., Ferreras J.M., Citores L.;  
 RT "Molecular cloning of ebulin 1."  
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: A7400822; CAC33178.1; --  
 DR HSSP: P02879; ZAAT.  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KM Glycosidase; Hydroxylase; Signal; Toxin.  
 FT SIGNAL 1 25  
 FT CHAIN 26 298 POTENTIAL.  
 FT CHAIN 299 564 EBULIN L A-CHAIN.  
 FT CHAIN 299 564 EBULIN L B-CHAIN.

SQ SEQUENCE 564 AA; 62694 MW; 8261681A6DB55CB8 CRC64;  
 Query Match 33.5%; Score 323.5; DB 10; Length 564;  
 Best Local Similarity 40.0%; Pred. No. 2.6e-22;  
 Matches 76; Conservative 38; Mismatches 51; Indels 25; Gaps 6;  
 QY 6 YPIINFTTAGATVOSTYNTFIRAVR-----LTVLPNRVGLPINRFLIVELSNH 55  
 DB 28 YPSVSNLAGAKSTYDEFLKRLDRVATGCTEVNGLPLRRESEYQVKNRFLVRLTNV 87  
 QY 56 AELSVTLALDYTNNAVYVGRAGNSAYFFHPNODDAEAT--TLFTDYQNRRTFAFGN 113  
 DB 88 NGDYTSVAVVNTLVAFSANGSYFF-----KDALEQKSNLFLGT--TQHTLSFTGNV 141  
 QY 114 DRLEQLGNLRNIELGNPLLEAISAALYYSTGGTQPLTLARSPFLICIONISEARFOY 173  
 DB 142 DNLETAAGTRRESLELGNPLDGAITSMT--DGG-----VANSLLVLIOMVEAARFK 194  
 QY 174 IEGEMRTIR 183  
 DB 195 IEQEVRRSLQ 204

RESULT 11  
 Q06076 PRELIMINARY; PRT; 528 AA.  
 ID Q06076;  
 AC Q06076;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Abrin-d (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).  
 OS Abrus precatorius (Indian licorice) (Crab's eye).  
 CC Eukaryota; Viridiplantae; Magnoliophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eucosids; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 OX NCBI\_TaxID=3816;  
 RN NCB1\_TaxID=3816;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93132798; PubMed=8421313;  
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;  
 RT "Primary structure of three distinct isoabrin determined by cDNA  
 RT sequencing: conservation and significance."  
 RL J. Mol. Biol. 229:263-267(1993).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL: M98346; AAA32626.1; --  
 DR HSSP: P11140; IABR.  
 DR InterPro: IPR000772; Ricin\_B\_lectin.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: PR00396; SHIGARICIN.  
 DR SMART: SM00458; RICIN; 2.  
 DR PROSITE: PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KM Hydroxylase; Toxin.  
 FT NON\_TER 1 1  
 FT NON\_TER 528 528  
 SQ SEQUENCE 528 AA; 58870 MW; 62EDX2FB8FBE60F8 CRC64;  
 Query Match 33.0%; Score 318.5; DB 10; Length 528;  
 Best Local Similarity 43.3%; Pred. No. 7.3e-22;  
 Matches 81; Conservative 22; Mismatches 69; Indels 15; Gaps 4;  
 QY 5 QYPIINFTTAGATVOSTYNTFIRAVRGLT-----VLPNRVGLPINRFLIVELSNH 56  
 DB 1 QDQVKEFTTEGATSQSKOPELEALRQLTGGLHDIPDPPTVEERNRYIVELSNSE 60  
 QY 57 ELSTVTLALDYTNNAVYVGRAGNSAYFFHPNODDAEATLTLFTDYQNRRTFAFGNVDRL 116  
 DB 61 RESIEVGIDVTNAVYVGRAGNSAYFFHPNODDAEATLTLFTDYQNRRTFAFGNVDRL 116





